Search for dark matter produced in association with bottom quarks with the ATLAS detector at the Large Hadron Collider in pp collisions at $\sqrt{s} = 13$ TeV

Edward James Thorpe

School of Physics and Astronomy

Queen Mary University of London

Submitted in partial fulfillment of the requirements of

Doctor of Philosophy

September 2020

Declaration

I, Edward James Thorpe, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date:

Details of collaboration and publications:

This work was supported by the Science and Technology Facilities Council

Edward James Thorpe
September 2020

Acknowledgements

And I would like to thank Marcella Bona for her supervision and guidance throughout my PhD. For their technical supervision, I would like to thank Alison Eliot, Calum MacDonald and John Anders. I would like to thank the bb+MET team with whom I collaborated endlessly. I would like to thank Emmy Gabriel and Lewis Wilkins for the love and support throughout my PHD. Lastly I would like to thank my Family and particularly my brother Matt who paved the way for me to do this.

Abstract

This thesis documents the results of a search for weakly interacting massive dark matter particles produced in association with bottom quarks. Final states containing bottom quarks and missing transverse momentum are considered. This search uses 139.0 fb^{-1} of proton-proton collision data as recorded by the ATLAS experiment at the LHC at CERN at $\sqrt{s} = 13$ TeV between 2015 and 2018. The results are interpreted in the framework of simplified models of colour-neutral scalar and pseudo-scalar mediators. This search sets expected lower limits on the production cross section of between 10 and 40 times the nominal for mediator masses between 10 and 100 GeV assuming a dark matter mass of 1 GeV and a unitary coupling.

Also presented is a study of the rates and efficiencies of the Level 1 Calorimeter Trigger of the ATLAS detector. This was done by simulating the trigger hardware in an offline analysis to be able to manipulate online parameters and study the effect this has on the simulated hardware response. Studies are shown looking at a variety of pile-up regimes and run conditions using data from 2016, 2017 and 2018 proton-proton collisions recorded by the ATLAS experiment. Signal efficiencies are measured using $Z \to e^+e^-$ tag-and-probe samples, while rates are obtained from enhanced minimum bias samples.

Table of contents

List of figures

List of tables

Chapter 1

Theoretical Review

1.1 The Standard Model

The combined and cumulative efforts of physicists over the this century and the last have contributed to the formulation of the Standard Model (SM) of particle physics [1]. The SM can be described as a set of equations containing everything that has so far been theoretically explained about the fundamental nature of reality at the quantum scale. The mathematical framework used to formulate this equation is quantum field theory (QFT), which is a combination of classical field theory, special relativity and quantum mechanics. The SM relies on mathematical symmetries and is described by the local gauge symmetry group $SU(3) \times SU(2) \times U(1)$. The $SU(3)$ component here pertains to the strong nuclear force, which accounts for the interactions between quarks and gluons, while the $SU(2) \times U(1)$ symmetry group corresponds to the combination of the electromagnetic (EM) and the weak nuclear forces unified in the electroweak theory, which accounts for the interactions between photons and vector bosons with quarks.

Although the SM is the most complete, experimentally verified quantum mechanical framework currently available, there are still a plethora of phenomena that have been observed in nature that can not be explained by the SM in its present formulation. In this chapter an overview of the SM is presented as well as a discussion about some of its theoretical and phenomenological shortcomings with particular focus on dark matter (DM) and how it could potentially be observed in a collider experiment.

1.1.1 Particles

There are two types of fundamental particle that are described in the SM, the first of which are the fermions. These are particles that have half-integer spin and obey Fermi-Dirac statistics

which describes their energy distribution. The nature of this distribution is dictated by the fact that no two fermions can occupy the same quantum state as dictated by the Pauli exclusion principle. The other type of SM particles, the bosons, are described by Bose-Einstein (BE) statistics which do allow particles to occupy the same state. These particles have integer spin and the spin-1 bosons function as the mediators for the three fundamental forces described by the SM.

Fermions

Summarised in Table 1.1 are the 12 elementary fermions. The fermion category can be further divided into quarks and leptons. Leptons come in three flavours corresponding to three generations and each generation has a charged lepton/neutrino pair. In order from first to third generation there is the electron and electron-neutrino, the muon and muon-neutrino and the tau and tau-neutrino. The mass of the charged leptons increases with each generation and each charged lepton has a charge -e, where e is taken as the absolute value of the charge of the electron.

Generation	Type	Name	Symbol	Charge (Q)	Colour	Mass
1st	quark	up	u	$+\frac{2}{3}e$	r,g,b	$2.16^{+0.49}_{-0.26}$ MeV
	quark	down	d	$-\frac{1}{3}e$	r,g,b	$4.67^{+0.48}_{-0.17}$ MeV
	lepton	electron	e	-e	-	$0.510998928 \pm 0.000000011$ MeV
	lepton	electron-neutrino	ν_e	0	-	< 2 eV (CL = 95 %)
2nd	quark	charm	c	$+\frac{2}{3}$e	r,g,b	1.27 ± 0.02 GeV
	quark	strange	s	$-\frac{1}{3}e$	r,g,b	93^{+11}_{-5} MeV
	lepton	muon	μ	-e	-	$105.6583715 \pm 0.0000035$ MeV
	lepton	muon-neutrino	ν_μ	0	-	< 0.19 eV (CL = 90 %)
3rd	quark	top	t	$+\frac{2}{3}e$	r,g,b	172.9 ± 0.4 GeV
	quark	bottom	b	$-\frac{1}{3}e$	r,g,b	$4.18^{+0.03}_{-0.02}$ GeV
	lepton	tau	τ	-e	-	1776.82 ± 0.16 MeV
	lepton	tau-neutrino	ν_τ	0	-	< 18.2 MeV (CL = 95 %)

Table 1.1 Summary of the fermions. Q is the electric charge of the particle and e is the absolute value of the charge of the electron. All masses are taken from the particle data group (PDG) and the top mass given here is taken from the results of direct measurements [2].

Quarks are fermions that can either be up-type or down-type. The up-type quarks are, in order of increasing mass and generation, the up quark (u), the charm quark (c) and the top quark (t). The top quark is the heaviest known fundamental particle weighing almost as much as a gold atom. These quarks all have a charge of $\frac{2}{3}e$. Similarly, the down-type quarks, with a charge of $-\frac{1}{3}e$, are the down quark (d), the strange quark (s) and the bottom quark (b).

Quarks have colour charge, which means that they can interact via the strong force. The colour charge of a quark can either be red (r), green (g) or blue (b) and the colour charge of an anti-quark can be either antired ($\bar{r}$), antigreen ($\bar{g}$) or antiblue ($\bar{b}$). Quarks are not observed in nature as lone particles and are generally confined in colourless triplets (rbg or $r\bar{b}\bar{g}$) or doublets ($r\bar{r}$,$b\bar{b}$ or $g\bar{g}$). Particles formed of doublets are called mesons and particles formed from triplets are called baryons.

Leptons, on the other hand, do not possess colour charge and so can only interact with each other via the EM force and the weak force. The neutrinos in the SM are only accounted for as massless particles, however experimental observations of neutrino oscillations mean that neutrinos must have some mass, although the experimentally established upper limits on these masses is very very small (mass scale $<\sim$ 1.1 eV [3]).

Bosons

The fundamental bosons, as detailed in Table 1.2, are integer spin (either 0 or 1) particles which obey BE statistics. Spin-1 bosons mediate the forces described by the SM. The photon is the mediator of arguably the most tangible of these forces, EM. Photons are massless and have infinite range. Electric charge is conserved in all known particle interactions making this a fundamental symmetry of nature.

Name	Symbol	Charge (Q)	Colour	Spin	Mass
photon	γ	0	-	1	0
W-bosons	$W^{\pm}$	$\pm$e	-	1	80.379 ± 0.012 GeV
Z-boson	Z	0	-	1	91.1876 ± 0.0021 GeV
gluon	g	0	8 combinations	1	0
Higgs	H	0	-	0	125.10 ± 0.14 GeV

Table 1.2 Summary of the bosons. Q is the electric charge of the particle, e is the absolute value of the charge of the electron and spin is measured in units of the reduced Planck constant $\hbar$. All masses are taken from the PDG [2].

The $W^{\pm}$ and Z bosons bosons are the mediators of the weak force. Unlike the photon, the $W^{\pm}$ and Z bosons do have mass and are finite in range. This is the force which is partly responsible for the structure of the nucleus of an atom. The rest of this structure is accounted for by the strong force. This is mediated by the gluon which is massless, however the strong force does have a finite range. The finite range of the strong force is due to a property of the theory called 'asymptotic freedom', which results in the force becoming stronger at larger

distances. This property of the strong force is the reason why quarks and gluons are bound within hadrons.

The SM has evolved as experiments have shed light on some of its shortcomings and previous iterations of the SM only predicted these spin-1 bosons and also left no mechanism for them to acquire mass. However, the mass of the $W^{\pm}$ and Z bosons were eventually measured and therefore the theory could not have been complete. A proposed solution to the problem of the masses of these vector bosons was the introduction of a new scalar (spin-0) field. The introduction of this new field also required a massive scalar particle. The new particle is now known as the Higgs boson and it was discovered by the combined efforts of the ATLAS and CMS experiments at the Large Hadron Collider at CERN in 2012 [4][5].

1.1.2 Symmetries and Lagrangians

Lagrangian terms in the SM take the form

$$\mathscr{L} = \mathscr{L}\left(\psi, \partial_\mu \psi\right), \tag{1.1}$$

where ψ is a fermion field and $\partial_\mu \psi$ is the partial four-derivative of that field in 4-dimensional space. If a field is invariant under the phase transformation

$$\psi \to e^{-i\alpha}\psi, \tag{1.2}$$

where α is a phase, then this transformation is a global symmetry of the field in nature. The SM only requires symmetries to be local which means that the field, ψ, will remain invariant under a spatially dependent phase transformation, which can be denoted as

$$\psi(x) \to e^{-i\alpha(x)}\psi(x). \tag{1.3}$$

The kinetic terms of Lagrangians are written as a function of $\partial_\mu \psi$, which would not remain invariant under the transformation given in Equation 1.3. The extra terms arising from this transformation can be eliminated by defining a new 'covariant derivative' of the form

$$D_\mu = \partial_\mu - ieA_\mu, \tag{1.4}$$

where A_μ is a vector field which transforms as

$$A'_\mu = A_\mu - \frac{1}{e}\partial_\mu \alpha(x). \tag{1.5}$$

Defining A_μ and D_μ in this way allows for the covariant derivative of the field to transform in the same way as the field itself under this local transformation. The additional terms generated by the partial derivative are absorbed by the transformation of A_μ in the covariant derivative and therefore kinetic terms of the form

$$\overline{\psi} D_\mu \psi \tag{1.6}$$

will be invariant under this local transformation.

These transformations form the basis of our formulation of the SM in this chapter. In the next section a specific implementation of this will be explored.

1.1.3 QED

In quantum electrodynamics (QED) the Lagrangian can be constructed starting with the Dirac Lagrangian given by

$$\mathscr{L}_{\text{Dirac}} = i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi, \tag{1.7}$$

where γ^μ are the Dirac matrices $\left\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\right\}$. Under the local phase transformation, as described in Equation 1.3, this can be expanded to

$$\begin{aligned}
\mathscr{L}'_{\text{Dirac}} &= i\overline{\psi}'\gamma^\mu \partial_\mu \psi' - m\overline{\psi}'\psi' \\
&= i\overline{\psi} e^{-i\alpha(x)}\gamma^\mu \partial_\mu \left[e^{i\alpha(x)}\psi\right] - m\overline{\psi} e^{-i\alpha(x)} e^{i\alpha(x)}\psi \\
&= i\overline{\psi} e^{-i\alpha(x)}\gamma^\mu \left(e^{i\alpha(x)}\partial_\mu \psi + ie^{i\alpha(x)}\psi \partial_\mu [\alpha(x)]\right) - m\overline{\psi}\psi \\
&= i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi + i\overline{\psi}\gamma^\mu \psi \partial_\mu [\alpha(x),]
\end{aligned} \tag{1.8}$$

where $\alpha(x)$ is the phase of the transformation.

Using the covariant derivative, as defined in Section 1.1.2, the Lagrangian can be written as

$$\begin{aligned}
\mathscr{L}'_{\text{Dirac}} &= i\overline{\psi}'\gamma^\mu D'_\mu \psi' - m\overline{\psi}'\psi' \\
&= i\overline{\psi} e^{-i\alpha(x)}\gamma^\mu \left(\partial_\mu + ieA_\mu - i\partial_\mu \alpha(x)\right)\left[e^{i\alpha(x)}\psi\right] - m\overline{\psi} e^{-i\alpha(x)} e^{i\alpha(x)}\psi \\
&= i\overline{\psi} e^{-i\alpha(x)}\gamma^\mu \Big(e^{i\alpha(x)}\partial_\mu \psi + ieA_\mu e^{i\alpha(x)}\psi \\
&\quad -ie^{i\alpha(x)}\psi \partial_\mu [\alpha(x)] + ie^{i\alpha(x)}\psi \partial_\mu [\alpha(x)]\Big) - m\overline{\psi}\psi \\
&= i\overline{\psi}\gamma^\mu D_\mu \psi - m\overline{\psi}\psi,
\end{aligned} \tag{1.9}$$

where A_μ corresponds to the photon field. The kinetic term for the QED Lagrangian is defined using

$$F_{\mu\nu} = ie\left(\left[\partial_\mu A_\nu\right] - \left[\partial_\nu A_\mu\right]\right), \tag{1.10}$$

where $F_{\mu\nu}$ is the electromagnetic field tensor. The complete QED Lagrangian can then be written as

$$\mathscr{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi}\ \not{D}\psi - m\overline{\psi}\psi, \tag{1.11}$$

where $\not{D} = \gamma^\mu D_\mu$. This Lagrangian is now, by construction, totally gauge invariant. An interesting exercise here is to try to add in a Dirac mass term for A_μ. Looking only at the mass term one can write

$$\mathscr{L} = m^2 A_\mu A^\mu, \tag{1.12}$$

which transforms via a local gauge transformation as follows

$$\begin{aligned}\mathscr{L}' &= m^2 A'_\mu A'^\mu \\ &= m^2\left(A_\mu - \frac{1}{e}\partial_\mu\alpha(x)\right)\left(A^\mu - \frac{1}{e}\partial^\mu\alpha(x)\right) \\ &\neq m^2 A_\mu A^\mu.\end{aligned} \tag{1.13}$$

By introducing a Dirac mass term into the QED Lagrangian we have made it no longer gauge invariant, therefore this cannot be used as a mechanism by which particles can acquire mass.

1.1.4 QCD

Quantum chromodynamics (QCD) is derived using the $SU(3)$ symmetry group and describes the strong interactions between quarks and gluons. The Lagrangian can be built in a similar way to the QED Lagrangian, but the electric field tensor term $F^{\mu\nu}$ is replaced by the gluon field tensor term $G^\alpha_{\mu\nu}$, which is defined later. In QCD, quarks transform as colour triplets that can be represented as

$$\psi_f(x) = \begin{pmatrix} \psi^r_f(x) \\ \psi^g_f(x) \\ \psi^b_f(x) \end{pmatrix}, \tag{1.14}$$

where f is the flavour index (u, d, s, etc..) and r, g, b are the colour indices.

The gluon field tensor is then used to construct the QCD Lagrangian of the form

$$\mathscr{L}_{QCD} = \sum_f \overline{\psi}_g^b \left(i \not{D}_{QCD} - m_f \right) \psi_f^b + \frac{1}{4} G_{\mu\nu}^{\alpha} G^{\alpha\mu\nu}, \tag{1.15}$$

where b is the colour charge given by r, g, b for quarks and $\overline{r}$, $\overline{g}$, $\overline{b}$ for anti-quarks and $\alpha \in [1,\ldots,8]$). This is because there are $N_c^2 - 1$ massless gluons, where N_c is the number of colours. Here we have also introduced a new covariant derivative. This can be written in the form

$$\not{D}_{QCD} = \gamma^{\mu} \left(\partial_{\mu} + i g_s T^{\alpha} A_{\mu}^{\alpha} \right), \tag{1.16}$$

where, similarly to QED, A_{μ}^{α} represent the gluon gauge fields. g_s is the gauge coupling term which is analogous to the electric charge, e. T^{α} are the generators of the $SU(3)$ symmetry group, which obey the Lie algebra

$$\left[T^i, T^j \right] = i f^{ijk} T^k, \tag{1.17}$$

where the structure algebra term f^{ijk} is introduced which is totally anti-symmetric under $SU(3)$ transformations.

The gluon field tensor can be written as

$$G_{\mu\nu}^{\alpha} = \partial_{\mu} A_{\nu}^{\alpha} - \partial_{\mu} A_{\nu}^{\alpha} - g_s f^{ijk} A_{\mu}^{j} A_{\nu}^{k}. \tag{1.18}$$

The terms in the field tensor which have a dependence on g_s lead to the self coupling of the QCD mediators, which gives rise to some of the more important phenomenological differences between QCD and QED. In QED the strength of interaction between two particles falls off in the familiar way as $\frac{1}{r^2}$, whilst the strong force exhibits asymptotic freedom meaning that the strength of the force decreases at smaller distance scales. This leads to quark confinement as when quarks move further apart it becomes energetically favourable to produce a quark and an anti-quark. This means that quarks are only ever found as hadrons in nature at energies below the QCD scale Λ_{QCD}. This scale is usually approximated to be close to the pion mass (~ 200 MeV) as the pion is the lightest of the hadrons. The experimental consequence of this for collider physics is that single quarks are not observed, but instead quarks are seen in cascades known as jets.

1.1.5 Electroweak Unification

The EM force and the weak force can be unified under a single combined symmetry group $SU(2)_L \times U(1)_Y$. Only components of fields with left-handed chirality (or "handedness") transform under $SU(2)_L$, which is the symmetry group of the weak field. $U(1)_Y$ corresponds to the EM symmetry group, however we are now considering the force to be acting upon hypercharge defined as

$$Y = 2\left(Q - T_3\right), \tag{1.19}$$

where Q is the electric charge of the particle and T_3 is the third component of weak isospin, which is a quantum number that must be conserved under the weak interaction. Hypercharge is a conserved quantity in the SM. The combined electroweak Lagrangian can be written as

$$\mathscr{L}_{EW} = \sum_f \overline{\psi}_f i \not{D}_{EW} \psi_f - \frac{1}{4} W_i^{\mu\nu} W^i_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}, \tag{1.20}$$

where we have the gauge fields $W^{i=1,2,3}_{\mu\nu}$, which transform under $SU(2)_L$, and $B_{\mu\nu}$ which transforms under $U(1)_Y$. As was done for the QED and QCD Lagrangians, one can define the covariant derivative as follows:

$$\not{D}_{EW} = \gamma^\mu \left(\partial_\mu + ig' \frac{Y}{2} B_\mu + ig \frac{\tau^i_L}{2} W^i_\mu \right), \tag{1.21}$$

where τ^i_L are the generators of the $SU(2)$ symmetry group and g and g' are coupling constants. The W^α and B terms describe the self-interaction of the gauge bosons. The kinetic term is given by

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g_{\mathrm{W}} \varepsilon^{ijk} W^j_\mu W^k_\nu, \tag{1.22}$$

where ε^{ijk} are the structure constants of the $SU(2)$ symmetry group and the self interaction term is proportional to the coupling, g_{W}. The physical interpretation of the gauge bosons here are the mixed states

$$\begin{aligned} W^\pm_\mu &= \tfrac{1}{\sqrt{2}} \left(W^1_\mu \mp i W^2_\mu \right), \\ Z_\mu &= B^0_\mu \cos(\theta_W) + W^3_\mu \sin(\theta_W), \\ A_\mu &= -B^0_\mu \sin(\theta_W) + W^3_\mu \cos(\theta_W), \end{aligned} \tag{1.23}$$

where θ_W is some mixing angle. $W^\pm_\mu$ is a charged massless state that corresponds to the $W^\pm$ bosons. Z_μ and A_μ are neutral massless states which correspond to the Z boson and

the photon respectively. We know from experiments that the $W^{\pm}$ and Z bosons are indeed massive, however there is no term in the electroweak Lagrangian to account for this. In order to generate the masses of these bosons it is necessary to break electroweak symmetry and introduce the Higgs boson.

1.1.6 Spontaneous Symmetry Breaking and the Higgs Mechanism

Until now the SM has no term to account for the masses of the $W^{\pm}$ and Z bosons. In order to derive these, it is necessary to break electroweak symmetry. The mechanism to do this was discovered by Higgs, Brout and Englert [6–8]. If we take ϕ to transform under $SU(2)$ symmetry with weak hypercharge of $Y = +1/2$, then the gauge transformation of ϕ can be written as

$$\phi \rightarrow e^{i\alpha}\phi. \tag{1.24}$$

where α is a constant. The covariant derivative is given by

$$D_\mu \phi = \left(\partial_\mu + ig\tau^a W_\mu^a + i\frac{Y}{2}g' B_\mu\right)\phi. \tag{1.25}$$

The most general way that this can then be added into the SM Lagrangian is by defining

$$\mathscr{L}_{\text{Higgs}} = \left(D_\mu \phi\right)^2 + \mu^2 \phi^\dagger \phi - \lambda\left(\phi^\dagger \phi\right)^2 \tag{1.26}$$

with a kinetic term, a mass term and a self interaction term respectively. λ is the coupling strength of the self interaction of this field. By defining the vacuum expectation value of the Higgs field to be v, this field can be parameterised in terms of v and a fluctuating field $h(x)$ with an expectation value centered on 0 as follows:

$$\phi(x) = U(x)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.27}$$

where $U(x)$ is a gauge transformation in $SU(2)$ that is necessary to keep this spinor in the most general form possible. The vacuum expectation value can be written as a function of μ and λ by substituting it into the most general $SU(2) \times U(1)$ scalar potential which can be written as

$$V(\phi) = \mu^2\left(\phi^\dagger \phi\right) - \lambda\left(\phi^\dagger \phi\right)^2 \tag{1.28}$$

and setting the derivative equal to zero. This yields the equation

$$v = \left(\frac{\mu^2}{\lambda}\right)^{1/2}. \tag{1.29}$$

Using Equation 1.27 and Equation 1.26 and looking only at the terms with a dependence on h it is possible to rewrite the Higgs Lagrangian in the form

$$\mathscr{L}_{\text{Higgs}} = \frac{1}{2}\left(\partial_\mu h\right)^2 - \mu^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4. \tag{1.30}$$

Here we have a kinetic term, a mass term with mass $m_h = \sqrt{2}\mu = \sqrt{2\lambda}v$ and two self interaction terms. The particle predicted in this theory is the Higgs boson.

There are also additional terms that come from this expansion which have no dependence on h. Evaluating these and also expanding the Pauli matrices yields Lagrangian terms for the vector bosons,

$$\mathscr{L}_V = \frac{1}{2}\frac{v^2}{4}\left[g^2\left(W_\mu^1\right)^2 + g^2\left(W_\mu^2\right)^2 + \left(-gW_\mu^3 + g'B_\mu\right)^2\right]. \tag{1.31}$$

Here we finally have the mass terms for the vector bosons. One can redefine the gauge fields as

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \text{ with mass } m_W = g\frac{v}{2} \\ Z_\mu^0 &= \frac{1}{\sqrt{g^2+g'^2}}\left(gW_\mu^3 - g'B_\mu\right) \text{ with mass } m_Z = \frac{v}{2}\sqrt{g^2+g'^2}. \\ A_\mu^0 &= \frac{1}{\sqrt{g^2+g'^2}}\left(g'W_\mu^3 + gB_\mu\right) \text{ with mass } m_A = 0 \end{aligned} \tag{1.32}$$

The final terms from the expansion of Equation 1.26, which describe the interaction between the Higgs and the vector bosons, are given by

$$\mathscr{L}_{VVh,VVhh} = \left[m_W^2 W^{\mu+}W_\mu^- + \frac{1}{2}m_Z^2 Z^\mu Z_\mu\right]\cdot\left(\frac{2h}{v} + \frac{h^2}{v^2}\right) \tag{1.33}$$

This describes a 3-point interaction between the Higgs and two vector bosons as well as a 4-point interaction between two Higgs bosons and two vector bosons.

The final SM particles that are yet to have their masses explained here are the fermions. This can be achieved by considering a Yukawa type mass term of the form

$$g_f \overline{\psi}_L \phi \psi_R, \tag{1.34}$$

where g_f is the coupling between the Higgs field and the fermions. After spontaneous symmetry breaking (i.e. setting $\phi = \sqrt{\frac{1}{2}}(v+h(x))$) this becomes

$$g_f \frac{v}{\sqrt{2}} \overline{\psi}_L \psi_R. \tag{1.35}$$

The mass term for the fermion is then

$$m_f = \frac{g_f v}{\sqrt{2}}. \tag{1.36}$$

With this we have described all the particles in the SM. This theory has been tested extensively in high energy physics experiments and has shown to be remarkably robust. The discovery of the Higgs boson by the ATLAS and CMS collaborations was announced on 4 July 2012 providing the confirmation of the existence of the last remaining fundamental particle predicted by the SM [4, 9]. A summary of the SM cross section measurements by the ATLAS experiment, with the associated theoretical predictions, is shown in Figure 1.1.

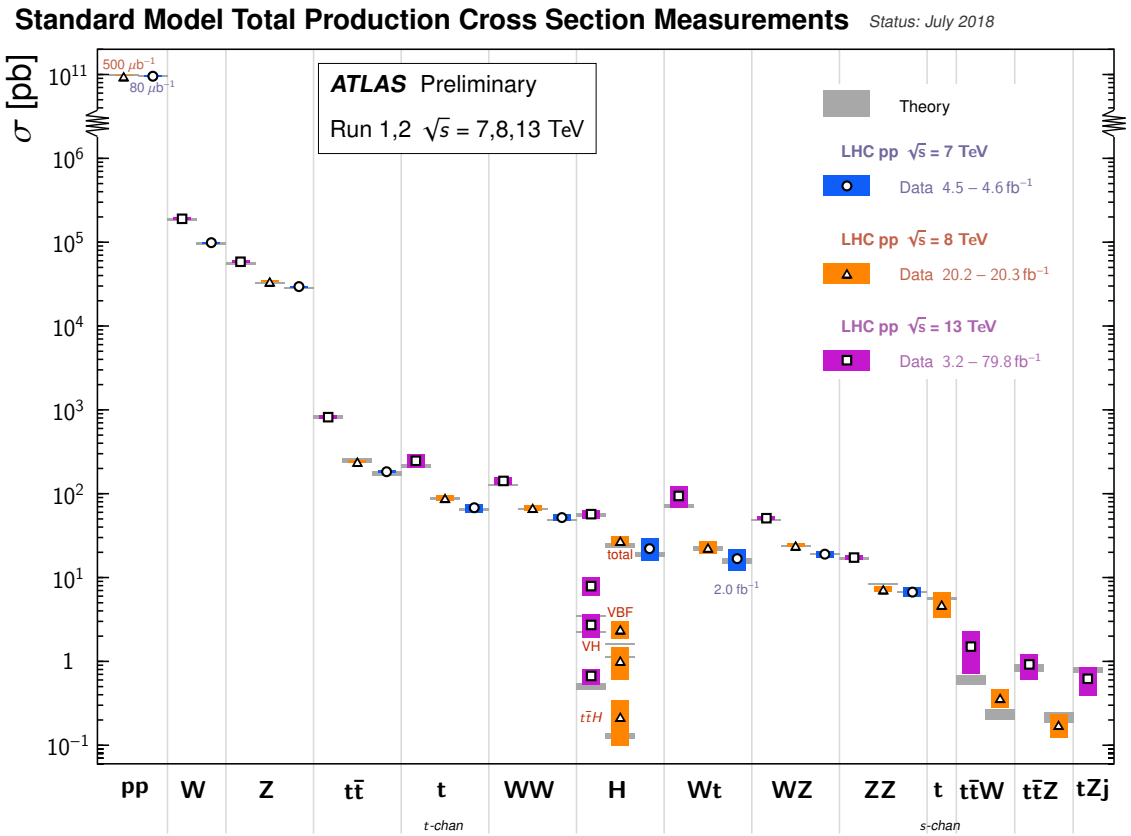


Fig. 1.1 Summary of SM cross section measurements, as measured by the ATLAS experiment as of July 2018, compared to theoretical predictions. [10]

1.2 Limitations of the Standard Model

The SM has been extraordinarily successful in predicting the existence of previously undiscovered particles and has remained incredibly robust against attempts to directly measure any deviation from it in its current form. However, The SM cannot be a complete theory of the natural physical world. There are known issues of the SM as well as a collection of physical phenomena that remain unaccounted for in this self-consistent theory. In this section, some of those issues will be explored.

1.2.1 The Hierarchy Problem

To understand the hierarchy problem one must first make the assumption that the SM is accurate up to some scale which we can define as Λ. We know that this can only be true at most up to the Planck scale, taken to be $M_P \sim 10^{19}$ GeV, which is the scale at which gravity becomes equally as potent as the other fundamental forces and therefore the SM must break down.

Figure 1.2 shows the loop correction to the Higgs mass arising from the self coupling of the Higgs. Integrating this loop diagram up until some energy scale Λ and only considering the highest order term in Λ yields

$$\Delta m_h^2 \phi^\dagger \phi \propto \lambda_{hhhh} \int^\Lambda d^4k \frac{1}{k^2 - m_h^2} \phi^\dagger \phi \sim \lambda_{hhhh} \Lambda^2 \phi^\dagger \phi, \tag{1.37}$$

where λ_{hhhh} is the self coupling constant and Δm_h is the quantum correction to the Higgs mass. If there is no new physics between the electroweak and the Planck scale then this correction will be very large ($\sim M_P$). This would mean that the corrections to the mass of the Higgs are dramatically larger than the mass of the Higgs itself ($\sim$125 GeV compared to $\sim 10^{19}$ GeV).

As well as the loop correction from the Higgs self-coupling, we can also consider the loop corrections from the other SM particles. The dominant correction will arise from a top quark loop which is shown in Figure 1.3. In a similar way to the loop correction that arises from the Higgs self-coupling, this can be written as

$$\Delta m_h^2 \phi^\dagger \phi \propto -\lambda_{t\bar{t}h}^2 \left(\int^\Lambda \frac{d^4k}{\not{k}\,\not{k}} \right) \phi^\dagger \phi \sim -\lambda_{t\bar{t}h}^2 \Lambda^2 \phi^\dagger \phi, \tag{1.38}$$

where $\lambda_{t\bar{t}h}$ is the Higgs coupling constant to $t\bar{t}$. A negative sign has been introduced here by virtue of the fact that we are now dealing with a fermion loop rather than a vector boson. If there is no new physics until the Planck scale then in order to get the measured

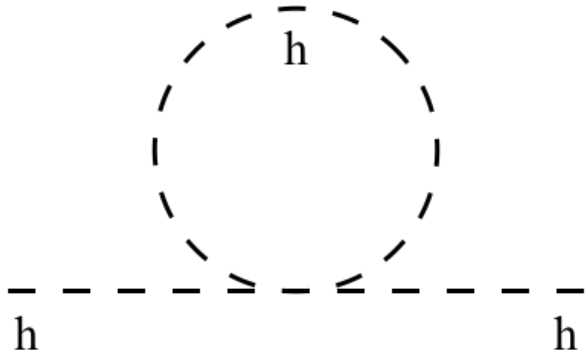


Fig. 1.2 Diagram corresponding to the loop correction to the mass of the Higgs boson that arises from its self-coupling.

Higgs boson mass, these loop corrections need to cancel to a precision 1 part in 10^{34}. This is regarded as being "unnatural" and is used as a motivation for the idea that new physics may well appear at energies much closer to the electroweak scale, which would negate the need for a large fine-tuning of parameters in the SM.

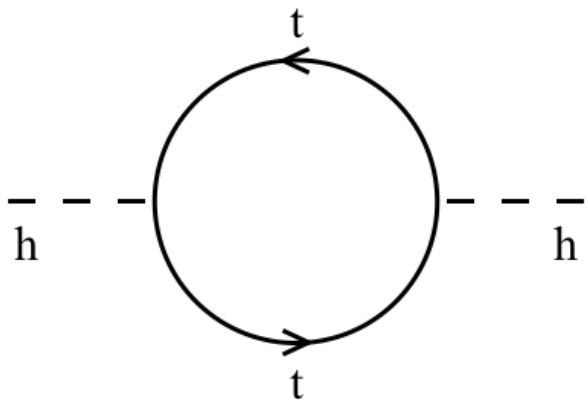


Fig. 1.3 The loop correction to the mass of the Higgs boson that arises from a top quark loop.

1.2.2 Gauge Coupling Unification

The idea of the unification of two forces is already a familiar one. Firstly there are the electric and magnetic forces which are unified under Maxwell's theory of EM and then there are the EM and the weak forces which have been consolidated into the electroweak force as seen in Section 1.1.5. In the context of the SM, unifying two forces means that one is able to describe them under the same symmetry group. The simplest way to do this for the SM is to assume that the current symmetries of $SU(3) \times SU(2) \times U(1)$ are a subset of an $SU(5)$

symmetry group. This would then be a grand unified theory (GUT) of physics. This theory would require $5^2 - 1$ gauge bosons of which 8 are accounted for by gluons in the SM and 4 are accounted for by the electroweak gauge bosons. One can then define the $X^{\pm\frac{4}{3}}$ and $Y^{\pm\frac{1}{3}}$ multiplets which account for the remaining 12.

There are several problems with this. One of which is that the running of the couplings in this model are predicted not to unify at GUT energy scales as shown in Figure 1.4. Also, although GUTs predict that the proton is not stable over very long time periods, current limits from water Cherenkov experiments place limits on the lifetime of this decay which are inconsistent with theoretically predicted limits from these simple GUTs.

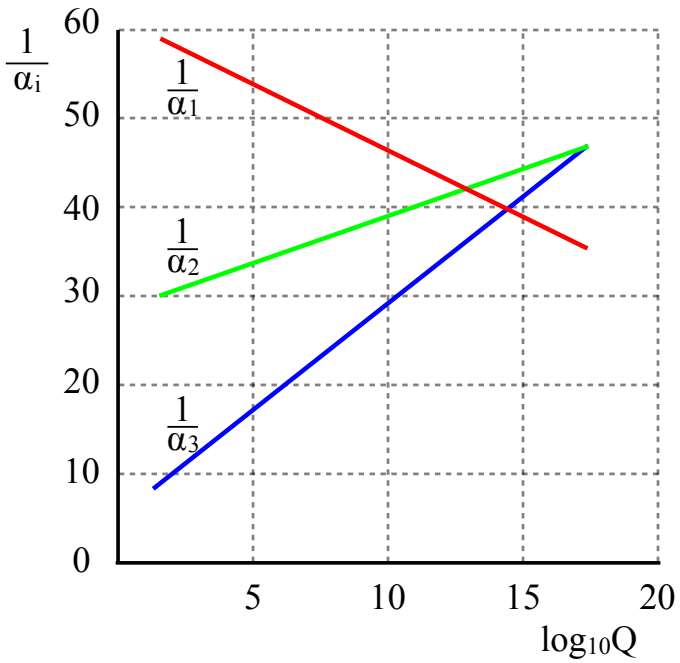


Fig. 1.4 The running of the gauge coupling in a simple GUT. On the x-axis is the log of the energy scale and on the y-axis is the inverse of the coupling strength. The EM coupling is in red, the weak coupling is in green and the strong coupling is in blue. [11]

1.2.3 CP Violation

Andrei Sakharov proposed three conditions that are necessary in order to explain why matter production dominated over antimatter in the very early universe. One of these conditions requires that CP symmetry is violated [12]. If it were not violated then equal numbers of left-handed baryons and right-handed anti baryons (and vice versa) would have been produced and would have annihilated.

CP violation is allowed in the SM. CP violation in the electroweak sector due to the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix has been experimentally observed [2]. The Pontecorvo–Maki–Nakagawa–Sakata (PMNS) lepton mixing matrix also contains a CP violating phase, which has not yet been experimentally measured. The current limits on this phase are consistent with zero [13]. If neutrinos are Majorana particles, meaning that they are their own anti-particles, then it is possible to get two more sources of CP violation via the PMNS matrix. It is also possible that CP violation could occur via the strong interaction. Attempted measurements of the electric dipole moment of the neutron (nEDM) have found this to be vanishingly small [14], however if CP violation were to manifest via this mechanism it should be finite (i.e. non-zero).

The measured sources of CP violation within the SM are not enough to account for the matter-antimatter asymmetry observed in the universe. This means that there must exist new physics with a source of CP violation that is able to account for this and so this can be considered to be another problem in our understanding of the universe that the SM is unable to explain.

1.2.4 Dark Matter

Although there exist other sources of evidence for BSM physics, the last one that will be addressed here is dark matter (DM). There exists a plethora of evidence for the existence of DM in the universe but the SM does not contain a viable candidate to explain it. The theoretical basis for DM will be explored in the next section as well as the sources of evidence for DM and candidate BSM particles that could account for it if it is particulate in nature.

1.3 Cosmology and Dark Matter

Very little is known about DM despite the fact that it constitutes $\sim$85% of the total matter density and around 26% of the energy density of the universe [15]. The precise nature of DM is not understood. Though there exist some possible candidates within the currently accepted formulations of the SM and general relativity, there also exist many candidates for DM in well-motivated extensions of the SM. These must be also be investigated. This section will cover some of the theoretical cosmological preliminaries, some of the sources for evidence of DM and how it may be possible to discover it in a collider experiment.

1.3.1 Cosmological Preliminaries

The three basic principles that are used to build the theoretical framework describing physics at the cosmological scale are as follows:

- The universe is homogeneous. Requiring homogeneity means that, on large enough scales, the universe is the same in all locations.
- The universe is isotropic. This is subtly different to requiring homogeneity as isotropy means that any measurement made of the universe will be independent of the direction of the measurement.
- The same laws of physics apply at all points in the universe. This means that the same laws of physics that apply on earth apply at every point in the universe. [16]

Numerous measurements of galactic redshift have shown that the universe is expanding [17]. The speed at which galaxies are moving away from each other can be related to the expansion of the universe with the very simple equation

$$v_{\text{rad}} = H_0 \cdot r, \tag{1.39}$$

where H_0 is the Hubble constant and r is the distance of the object being measured from the observer. The most recent results from the Hubble Space Telescope have measured this value to be 69.8 ± 1.9 kms^{-1}Mpc^{-1} [18].

A metric for this universe can be found by solving Einsteins field equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\frac{8\pi G_N}{c^4} T_{\mu\nu} - \Lambda g_{\mu\nu}, \tag{1.40}$$

where $g_{\mu\nu}$ is the metric tensor, $R_{\mu\nu}$ and R are the Ricci tensor and Ricci scalar respectively, G_N is Newton's gravitational constant and $T_{\mu\nu}$ is the stress-energy tensor. The metric is then

$$ds^2 = dt^2 - a^2(t) \left(\frac{dr^2}{1-kr^2} + r^2 \left(d\theta^2 + \sin\theta d\phi^2 \right) \right) = g_{\mu\nu} dx^\mu dx^\nu. \tag{1.41}$$

The choice $k = -1, 0, 1$ corresponds to the spatial curvature of the universe and θ, ϕ and r correspond to the usual angular and radial coordinates in the spherical coordinate system, and $a(t)$ is the scale factor of the universe at a time t. For this summary, the choice of $k = 0$ is made which corresponds to a flat universe. For this choice of k, the metric simply reduces

to that of Euclidean space. Substituting Equation 1.41 in Equation 1.40 yields the Friedmann equations

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 + \frac{kc^2}{a(t)^2} = \frac{8\pi G_N}{3} \cdot \rho_{\text{tot}}(t), \tag{1.42}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G_N}{3}\left(\rho_{\text{tot}}(t) + \frac{3p_{\text{tot}}(t)}{c^2}\right), \tag{1.43}$$

where p_{tot} is the total pressure density of the universe and ρ_{tot} is the total energy density of the universe. ρ_{tot} is defined as the sum of the matter density, ρ_{matter}, the radiation density, $\rho_{\text{radiation}}$, and the vacuum density, ρ_Λ, of the universe . The equation of state that defines the relationship between the total pressure and energy densities is

$$p_{\text{tot}}(t) = \omega(\rho)\rho_{\text{tot}}(t) \tag{1.44}$$

where $\omega(\rho)$ is the barotropic parameter. By solving the first Friedmann equation (1.42) with k=0, one can calculate the critical density by rearranging Equation 1.42

$$\rho_{\text{crit}}(t) \equiv \rho_{\text{tot}}(t)|_{k=0} = \frac{3H(t)^2}{8\pi G_N}, \tag{1.45}$$

where $H(t) = \frac{\dot{a}(t)}{a(t)}$. This relationship known as Hubble's law. This is defined such that at time t_0 the value of the Hubble constant is $H(t_0) = H_0$. The density parameter is defined as follows

$$\Omega_i(t) = \frac{\rho_i(t)}{\rho_{\text{crit}}(t)}. \tag{1.46}$$

The subscript i can be any of the contributions to the energy density of the universe (matter, radiation etc..). $\Omega_{\text{matter}}(t_0)$ has been measured to correspond to approximately 30% of the matter content of the universe by the WMAP and SDSS experiments [19]. Baryonic matter, on the other hand, accounts for only around 5% of the content of the universe [20]. This missing component of matter is known as DM. The following section will explore some of the sources of evidence for DM in the universe.

1.3.2 Observational Evidence for Dark Matter

Although to date there are no direct observations of DM interacting with SM particles via any of the SM forces, there are many independent sources of evidence for DM in the universe. The first of these observations comes from galactic rotation curves. These observations date

all the way back to 1884 when Lord Kelvin used the velocity distribution of stars orbiting the center of our galaxy in order to estimate the mass of the galaxy. The mass was considerably higher than the mass as calculated from the visible matter in the galaxy and so he concluded "many of our stars, perhaps a great majority of them, may be dark bodies".

Over the 20th century there have been many measurements of galactic rotation curves which have yielded results such as is shown in Figure 1.5. If there were only the visible matter observed in the Milky Way then this rotation curve would be expected to drop off towards the outer reaches of the galaxy. Instead what happens is that it remains relatively flat. It can then be hypothesised that this is due to there being a DM component to the makeup of the Milky Way.

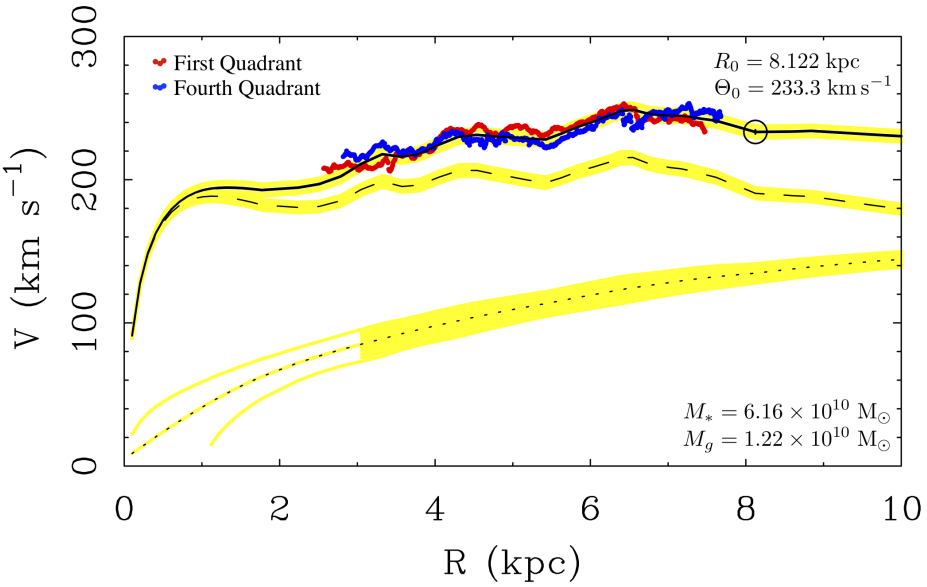


Fig. 1.5 Measurements of the velocities of stars around the center of the Milky Way. The solid black line shows a model that includes DM while the red and blue dots show points as measured from two quadrants of the galaxy. The dashed line shows the rotation curve of only the baryonic component and the dotted line shows the implied DM contribution. [21]

The measurement of the velocity of stars around the galactic center has been performed many times with some of the most recent results coming from the combined efforts of the Gaia and Hubble space telescopes. These collaborations calculated the mass of the Milky Way to be $1.28^{+0.97}_{-0.48} \times 10^{12} \mathrm{M}_\odot$, where $\mathrm{M}_\odot$ is the mass of the sun [21]. Of this, only around 5% is visible matter.

As well as this, gravitational lensing measurements have been able to observe DM. The general theory of relativity treats the force of gravity as massive objects curving space-time. Therefore very massive objects will bend light to a measurable degree. The ability to predict

the amount by which light should be bent around a massive object like a galaxy or a cluster of galaxies is one of many remarkable successes of general relativity.

One interesting example is the gravitational lensing observation made of two colliding clusters of galaxies known as the Bullet Cluster as shown in Figure 1.6. The gravitational lensing observation of the colliding clusters show that the center of mass of each one is displaced from the center of mass as calculated using only visible light [22–24]. It is not unanimously agreed upon in the scientific community that this spatial disparity is actually statistically significant [25]. If it is, then this seems to indicate that conventional theories of modified gravity would be unable to account for the disparity between these two measurements as theories like this mostly do not predict that the gravitational force that would act as though displaced from the center of mass of a given system. Therefore the Bullet Cluster observations further motivate the idea that DM may be particulate in nature and could be explained with an extension to the SM.

Fig. 1.6 The Bullet Cluster. The colours are false. In blue are the DM components of the colliding galaxies whilst in pink are the visible components. [26]

1.3.3 Dark Matter Candidates

There are plenty of well-motivated extensions to the SM that could possibly explain DM in the universe. Viable DM candidates are required to fulfil certain criteria. The first is that

they should be electrically neutral. If this were not the case then DM particles would scatter light which would be detectable via telescopes. DM particles must also be either stable or long lived on cosmological timescales in order to explain the abundance that is seen in the universe today. DM must be massive in order to explain the galactic rotation curve and gravitational lensing evidence. It must also be non-baryonic as the total energy density of the universe has been measured using the cosmic microwave background (CMB) and has been found to be inconsistent with the observed baryonic density in the universe [27]. Lastly, DM should be weakly interacting as it has been seen in astrophysical observations like those of the Bullet Cluster that DM is virtually collisionless [28]. The following is a non-exhaustive list of some of the most popular explanations for DM:

- Axions: these are pseudoscalar particles that could also solve the strong CP problem as laid out in Section 1.2 [29]. The mass of the axion has been constrained to be between around 10^{-3} to 10^{-5} eV [30].
- MACHOs: MACHO stands for 'massive astrophysical compact halo object'. These are objects which typically have a mass of around $\sim 0.08 M_{\odot}$ and are thought to be abundant in galactic halos. Due to the size of these objects they would emit very little radiation and so would not be seen by the currently available telescopes. Although they could potentially account for some of the DM, it is widely thought that these objects are not able to account for all of it [31].
- WIMPS: these are weakly interacting massive particles. These are hypothetical particles that could very weakly couple to SM particles and would be expected to have a mass at the GeV scale.

WIMPs

DM production could have taken place in the early universe when temperatures were much higher than the mass of the DM particles. DM production from SM particles would have taken place at a rate $\Lambda_A = n < \sigma_A v >$, where n is the number density of DM particle and v is their velocity and $< \sigma_A v >$ is the thermal average of the annihilation cross section. At this point, this rate would be equivalent to the rate of production of SM particles from DM. This would continue to occur until the rate of expansion exceeds the annihilation rate ($\Lambda_A > H(t)$).

After this DM freeze-out the relative DM density in the universe will remain constant and can be written as a function of the thermally-averaged annihilation cross section as

$$\Omega_{DM} \sim \frac{3 \times 10^{-27} \mathrm{cm}^3 \mathrm{s}^{-1}}{< \sigma_A v > H(t)^2}. \tag{1.47}$$

The value of $3 \times 10^{-27}\text{cm}^3$ is derived using the DM relic density $\Omega_\chi \simeq 0.22$.

Assuming that the DM is produced via s-channel exchange of an electroweak boson, the annihilation cross section is approximately $< \sigma_A v > \sim G_f^2 m_{DM}^2$, where G_f is the Fermi coupling constant, which denotes the strength of the interaction, and m_{DM} is the mass of the DM particle. Substituting this back into equation 1.47 yields the relation

$$\Omega_{DM} \sim (\frac{m_{DM}}{\text{GeV}})^2. \tag{1.48}$$

This means that the observed relic abundance is consistent with a DM mass around the GeV scale. This is known as the "WIMP miracle".

1.4 A Simplified Model Approach to Dark Matter Searches

Assuming that the relic abundance of DM is due to the existence of WIMPs, it must be the case that DM and SM particles are able to interact even if it is only very weakly. This means that it should be possible to create DM at pp collider experiments. From now on these experiments will be discussed in the context of the Large Hadron Collider (LHC) at CERN. There are three broad categories of theoretical framework that are employed at the LHC. The first of these are effective field theories (EFTs). EFTs don't contain granular information on the mediator particle. An example of an EFT in the context of DM is the pair production of two DM particles by two SM particles via a contact interaction as shown in Figure 1.7.

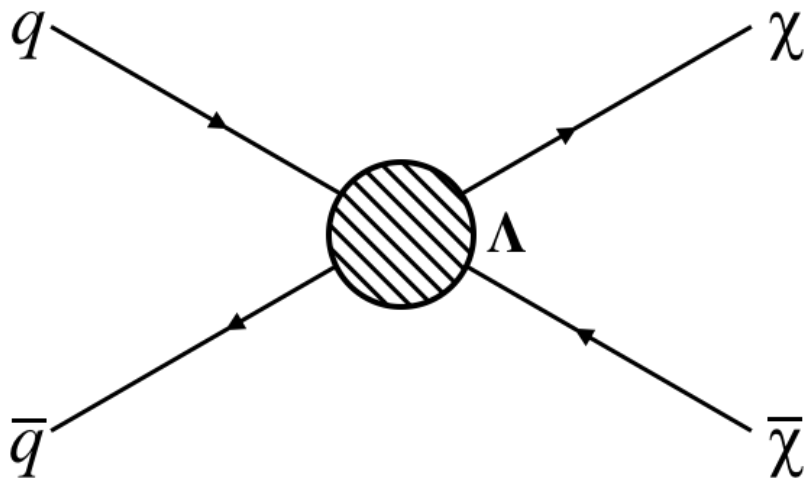


Fig. 1.7 An example of an EFT. Here a $q\bar{q}$ pair are creating a $\chi\bar{\chi}$ pair via some contact interaction at some energy scale Λ. This $\chi\bar{\chi}$ pair can be interpreted as a DM pair, but the methodology is more general than this.

The EFT approach was used during the 2011-2012 data taking period of the LHC (which is defined in Section 2.1). This approach has some shortcomings in the context of the LHC as the momentum transfer involved is of the order of the expected UV cut-off [32] and so for the 2015-2018 data taking periods of the LHC this approach is not widely used for DM searches as the center of mass energy of the LHC increase from 7 TeV to 13 TeV and so the assumption that the SM-DM mediator mass is comparable to the momentum transfer does not hold as true anymore.

Run 2 LHC DM searches commonly use 'simplified models'. These models state explicitly the spin and the mass of the mediator particle.

On the other end of the theoretical spectrum are UV complete models. Searches for simplified models can often be reinterpreted in the framework of a UV complete model.

This thesis will detail a search for the production of two b-quarks in association with a colour-neutral scalar (ϕ) or pseudoscalar (a) SM-DM mediator decaying into a pair of WIMP DM particles (χ) as shown in Figure 1.8. The coupling between the DM mediator and the SM model particles is highly restricted by precision flavour measurements, however assuming Minimal Flavour Violation (MFV) severely relaxes this constraint [33]. This allows us to keep the effects of CP violation, induced by new physics, small. Imposing MFV means that the interaction between SM particles and a new spin-0 neutral state is proportional to the mass of the fermions via Yukawa-type couplings. This motivates the choice to search for the production of these mediators in association with heavy flavour quarks. This thesis will focus only on the production in association with b-quarks. A choice is made to assume that the final DM particles ($\chi\bar{\chi}$) are 1 GeV in mass, which is heavily motivated from cosmology. The mass of SM-DM mediator particles will be in the range of 10 GeV to 500 GeV.

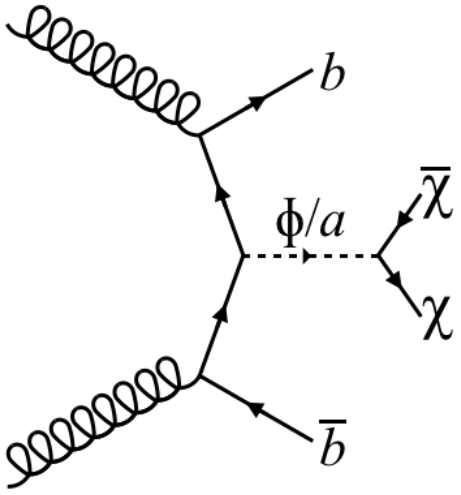


Fig. 1.8 *s*-channel tree level diagram of $b\bar{b}$ production in association with scalar (ϕ) or pseudoscalar (a) mediator decaying to a $\chi\bar{\chi}$ pair.

Chapter 2

The LHC and the ATLAS Detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is currently the most powerful particle collider in the world. It is a two-ring, superconducting proton-proton (pp) collider built in a tunnel of circumference 26.7 km with an average depth of 100 m. Construction of the LHC was completed in 2008 at the European Organisation for Nuclear Research (CERN) and runs underneath the border between France and Switzerland. The first data-taking periods were in 2011 and 2012, collectively known as 'Run 1'. During this time the LHC achieved center-of-mass energies ($\sqrt{s}$) of 7 TeV up until 2011 and then increasing to 8 TeV until the end of Run 1. The total integrated luminosity delivered during Run 1 was 5.46 fb^{-1}. The LHC has now finished taking data for Run 2, which took place between 2015 and 2018. During Run 2 the LHC was able to deliver 156 $fb^{-1}$of data to ATLAS at $\sqrt{s} = 13$ TeV recording a peak instantaneous luminosity, which is the number of collisions that can be produced per second and per cm^2, of 2.1×10^{34} cm^{2}s^{-1}.

The proton beam undergoes several stages of acceleration through multiple different accelerators prior to entering the LHC rings. Firstly, protons are extracted from a tank of hydrogen by stripping the hydrogen atoms of their electrons. These protons are then accelerated up to an energy of 50 MeV in the LINAC2 accelerator which is a linear accelerator. These are then fed into BOOSTER, which is the first of the three synchrotrons that the protons are accelerated by before reaching the LHC. Here the protons will be accelerated to 1 GeV, then up to 26 GeV in the Proton Synchrotron (PS) and then to 450 GeV in the Super Proton Synchrotron (SPS) which is sufficiently high energy to then enter the LHC where they will be accelerated to 6.5 GeV. Figure 2.1 shows an overview of the different accelerators at CERN as well as several detectors and experiments located at different points in the accelerator complex.

Fig. 2.1 A representation of the detectors and accelerators at CERN. The proton beam originates at LINAC2 eventually reaching maximum energy in the LHC. [34]

2.2 The ATLAS Detector

The ATLAS detector [35] is one of four main detectors at the LHC and one of two general purpose detectors. Covering a total solid angle of 4π, it is designed to cover as broad of a range of physics as possible including SM and possibly BSM processes. Figure 2.2 shows a diagram of the ATLAS detector and the various subsystems of the detector.

Fig. 2.2 Cut-away view of the ATLAS detector and each of the subsystems. [35]

The ATLAS coordinate system is, by convention, a right-handed coordinate system with its origin at the interaction point at the center of the detector. The z-axis is defined parallel to the proton beams, the x-axis points to the center of the LHC and the y-axis points upwards. The x-y plane is therefore transverse to the beam direction. For a particular event it is impossible to say what fraction of momentum the interacting parton is carrying, therefore the transverse system becomes particularly important as conservation of momentum can be used in this plane. Event kinematics are then described using ϕ, which is the angle in the x-y plane, and pseudo-rapidity. Pseudo-rapidity is the massless limit of rapidity, which is a Lorentz invariant measurement of angle relative to the beam direction. Rapidity is defined as

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right), \tag{2.1}$$

where E is the energy of the particle and p_z is the component of the particle's momentum in the direction of the z-axis. Pseudo-rapidity is then defined as

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{2.2}$$

where θ is the angle between the particle's vector momentum and the beam axis. Finally, the angular distance between two objects using these coordinates is given by

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{2.3}$$

The ATLAS detector consists of four main components:

- The magnet system: Powerful superconducting magnets are used to bend the tracks of charged particles so that their momenta can be measured.
- The inner detector: This is the innermost part of the detector and measures the tracks, momenta and charge of electrically charged particles that pass through it.
- The electromagnetic and hadronic calorimeters: Measures the energy of charged or neutral particles.
- The muon spectrometers: Identifies and measures the momenta of muons.

These are detailed in the following sections.

2.2.1 Inner Detector

The Inner Detector (ID) is the innermost sub-detector of ATLAS. The average time between bunch crossings in the ATLAS detector is 25 ns and during Run 2 the average number of interactions per bunch crossing ($<\mu>$) was 33.7. The ID is designed to be able to measure the tracks of the approximately 1000 charged particles produced with each bunch crossing. The three main sub-components of the ID, as shown in Figure 2.3, are the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT), all of which are immersed in the 2T magnetic field produced by the central solenoid magnet as detailed in Section 2.2.4.

Pixel Detector

The innermost sub-component of the ID, with an inner diameter of 101 mm, an outer diameter of 389 mm and length 650 of mm, is the pixel detector. It has a barrel and two end-cap

Fig. 2.3 The ATLAS Inner detector with major sub-components labelled [35].

components each of which has three layers of pixel sensors. The sensors are 50×400 μm^2 in size and are rotated slightly in the x-y plane such that they can then overlap and cover the entire range of $0 < \phi < 2\pi$. The pixel detector is made up of 1744 individual pixel sensors, each of which are 250 μm thick. Each sensor is made of an N-type silicon crystal which is designed to initially operate at a bias voltage of $\sim$150 V. However, to retain charge collection efficiency after years of radiation damage, these sensors were required to have an operating voltage of up to $\sim$600 V. The pixel detector can track the paths of charged particles in the pseudo-rapidity range $|\eta| < 2.5$. The pixel detector has a resolution of 14 μm in the $R-\phi$ plane and 115 μm in the z-direction.

As a part of the upgrade work between Run 1 and Run 2 an insertable B-layer (IBL) was put in place between the original pixel detector and the beam pipe. The inclusion of this additional high granularity pixel layer allowed for the distance of closest measurement to the beam to be reduced from $\sim$50 mm to $\sim$33 mm in the range $|\eta| < 2.9$. This enables a more precise measurement of primary and secondary vertices, which is particularly useful for the identification of b quarks [36]. The IBL has a resolution of 10 μm in the $R-\phi$ plane and 75 μm in the z-direction.

Fig. 2.4 Visualisation of the barrel components of the ID showing the pixel detector, the SCT and the TRT as well as the IBL. [37]

Semiconductor Tracker

The SCT employs a similar technology to the pixel detector. The main difference is that rather than using n-type silicon pixels, 15912 p-on-n silicon strips are used. Each sensor has a total of 768 strips which are 12 cm in length with a sensor thickness of 286 μm. Similarly to the pixel detector, the design voltage of the SCT is around $\sim$150 V but the operational voltage is required to be up to $\sim$350 V for Run 2, depending on the location of the particular strip, to account for radiation damage.

The barrel consists of 2112 modules. Each module is made up of four sensors rotated slightly in the x-y plane such that there are no gaps in the active area of the detector in ϕ. The end-cap is made up of a further 1976 modules, with three different geometries depending on the position of the particular module. Here each module is made of two sensors which are layered on top of each other and slightly rotated such as to cover as much solid angle as possible with active detecting components. The SCT has a resolution of 16 μm in the $R-\phi$ plane and 580 μm in the z-direction.

Transition Radiation Tracker

The basic technology used in the TRT component of the inner detector is the polyimide drift (or 'straw') tube. Made from a 25 μm thick polyimide film which is coated with 0.2 μm aluminium and then protected by a 5-6 μm thick graphite-polyimide layer, these drift tubes contain a mixture of 70% Xenon, 27% CO_2 and 3% O_2 gas [38]. The Xenon is for good X-ray absorption, while the combination of CO_2 and O_2 is used to increase the electron drift velocity and for photon-quenching. The potential difference is supplied by cathodes operating at $\sim$1530 V with a gain of 2.5×10^4. The tubes are embedded between 19 μm thick polypropylene fibres which serve as the transition material. A charged particle passing through the polypropylene will emit some EM radiation, the energy of which is logarithmically proportional to the Lorentz factor of the particle. This is particularly high for electrons and so the TRT, as well as providing additional tracking information, also provides identification information for electrons.

The barrel consists of 73 layers of tubes, interleaved with polypropylene fibres. Each tube runs the entire length of the barrel, with read outs at either end. They can therefore only yield tracking information in the $R-\phi$ plane, where R is defined as

$$R = \sqrt{\phi^2+\eta^2}. \tag{2.4}$$

The end-cap component is comprised of 160 straw planes interleaved with polypropylene foils for the electron identification.

Every charged track with $p_T > 0.5$ GeV and in the range $|\eta| < 2.0$ will traverse through at least 36 straws except for $0.8 < |\eta| < 1.0$ where the minimum number of straws traversed through by a charged particle is 22. The TRT has a resolution of 170 μm in the $R-\phi$ plane and 40-75 cm in the z-direction.

2.2.2 Calorimeters

Surrounding the inner detector and the solenoid magnet are the EM and hadronic calorimeters, which are shown in Figure 2.5. The EM calorimeter (ECal) is the innermost calorimeter component and it is designed to be able to detect and measure the energy of photons and electrons with fine granularity. The hadronic calorimeter (FCal), which has a coarser granularity, surrounds this and is designed to be able to detect and measure the energy of hadrons and jets of hadrons. Finally, there is the forward calorimeter (FCal), which is used to measure both hadronic and EM showers in the $3.1 < |\eta| < 4.9$ range. After noise subtraction, the energy resolution of a calorimeter is given by

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E[GeV]}} \oplus b, \tag{2.5}$$

where a is a stochastic term related to parton shower evolution and b is a constant term that accounts for inhomogeneities in the detector, calibration imperfections and leakage.

Containment of showers was an important design consideration for the calorimeter system. None of the energy from hadronic and EM showers should reach the muon systems (a phenomenon known as 'punch-through'). For the ECal the important quantity is the radiation length (X_0), which is the average length (cm) that a parton travels over which its energy will be reduced by a factor e. The HCal equivalent is the nuclear interaction length (λ_I), which is the mean distance travelled by a hadronic particle before undergoing an inelastic interaction with the detector material.

A common feature of the main components of the calorimeter system is an initial high density absorbing material. When a particle hits this material it produces a shower of secondary particles which are at much lower energy. These lower energy particles then continue through the active part of the detector where their energy is measured either via scintillation or ionisation.

Electromagnetic Calorimeter

The ECal consists of a barrel component covering the range $|\eta| < 1.475$ and two end-cap components covering $1.4 < |\eta| < 3.2$. The technology used to detect EM particles is the

Fig. 2.5 Cut away view of the ATLAS calorimeters showing the hadronic, EM and forward calorimeters. [35]

same in all components of this subsystem. Lead is used as the absorbing material and liquid argon (LAr) is used as the active material. These are then layered in an 'accordion shape' so as to provide the best possible ϕ coverage as shown in Figure 2.6. Three cryostats are used to keep the LAr at 88 K to keep it in liquid phase. The basic principle is that charged particles ionise LAr particles producing free electrons. These electrons then, under the influence of an applied electric field, drift towards an electrode where a signal can be read out.

Fig. 2.6 Photograph showing the accordion shape that the lead and LAr sheets are layered in. [39]

The barrel component is 0.53 m thick and 6.4 m long and was designed to be $> 22\ X_0$. The end caps are 0.63 m long and 1.77 m thick which corresponds to $> 24\ X_0$. The energy resolution of the ECal is $\frac{\sigma_E}{E[\text{GeV}]} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$.

Hadronic Calorimeter

The HCal uses the same principle as the ECal but with a different geometry, active material and absorbing material. It is located outside the ECal and consists of a tile calorimeter covering the range $|\eta| < 1.7$ and two LAr calorimeter end-cap components covering the range $1.5 < |\eta| < 3.2$. The barrel is divided into a central region ($|\eta| < 1.0$) and two extended barrel regions ($0.8 < |\eta| < 1.7$).

Both the central and extended tile calorimeters are divided into three layers. The central region has radiation lengths, from the innermost to the outermost layer, of 1.5, 4.1 and 1.8 λ_I. Similarly the three sections of the extended tile calorimeters have radiation lengths of 1.5, 2.6 and 3.3 λ_I respectively. Steel is used as the absorbing material, while plastic scintillating

tiles are used as the active material. The end-caps of the HCal use LAr as the active material and copper as the absorbing material. The resolution for the HCal is $\frac{\sigma_E}{E[\text{GeV}]} = \frac{50\%}{\sqrt{E}} \oplus 3\%$.

Forward Calorimeter

In the most forward region of the calorimeters is the FCal, which is in the range $3.1 < |\eta| < 4.9$ and is of maximum length 10 λ_I. It is built up of three layers, the first of which uses high density LAr and copper absorbing material. The next two layers, also use LAr as the active material, but use tungsten absorbers for hadronic measurements. The resolution of the FCal is up to $\frac{\sigma_E}{E[\text{GeV}]} = \frac{100\%}{\sqrt{E}} \oplus 10\%$.

2.2.3 Muon Spectrometer

As shown in Figure 2.7, the muon spectrometer (MS) is the outermost part of the ATLAS detector. It functions to identify muons as muons are the only detectable particle expected to reach this outer part of the detector. The MS also helps the ID accurately measure their momenta and functions as the first level trigger for muons. Muon momentum is determined by accurately measuring the tracks of muons as they pass through the detector and calculating their radius of curvature as they are bent by large superconducting magnets as discussed in section 2.2.4.

Monitored Drift Tubes (MDTs), which cover the range $|\eta| < 2.7$, are used for precision tracking in both the barrel and the end-cap regions. In the barrel, triggering is done using Resistive Plate Chambers (RPCs) for the range $|\eta| < 1.05$. The end-cap components are the Cathode Strip Chambers (CSCs), which are used for tracking, and the triggering is done using Thin Gap Chambers (TGCs). These cover the ranges $2.0 < |\eta| < 2.7$ and $1.05 < |\eta| < 2.7$ respectively.

Monitored Drift Tubes

The MDTs rely on muons ionising a pressurised gas of Ar (93%) and CO_2 (7%), which is contained within a tube $\sim$30 mm in diameter. The ionisation process creates free electrons which then drift in an applied electric field towards a central anode made of tungsten-rhenium where a signal can then be read out.

Cathode Strip Chambers

CSCs in the very high forward region ($2.0 < |\eta| < 2.7$) have a higher rate capability and time resolution than the MDTs. They use multi-wire proportional chambers, which rely on

Fig. 2.7 Cut away view of the ATLAS muon spectrometers highlighting the MDT, TGC, CSC and RPC components. [35]

ionisation to detect muons, and a cathode read out. The gas used for ionisation is made up of Ar, CO_2 and CF_4.

Resistive Plate Chambers

The RPC forms part of the first level muon triggers as well as contributing to the tracking of muons. The RPC relies on muons ionising a gas of mostly $C_2H_2F_4$. This component is formed of three layers perpendicular to the beam. The plates are made of a plastic laminate and an electric field is applied between the plates to allow a current to be detected by an anode.

Thin Gap Chambers

TGCs are multi-wire proportional chambers like CSCs. They rely on the same principles but instead use a gas which is a mixture of CO_2 and n-pentane. The TGCs have a very good time resolution, which is essential for triggering in the very high forward regions.

2.2.4 Magnet System

Accurate measurements of the momenta of particles is crucial to the performance of the ATLAS detector. The momentum of a charged particle can be determined by measuring the radius of curvature of that particle in a magnetic field. The momentum of a particle with charge q, in a magnetic field of strength B is given by

$$p = qBr, \tag{2.6}$$

where r is the radius of curvature of the particle. The ATLAS detector therefore requires powerful magnets such that the momenta of charged particles can be accurately determined. The magnet system is made up of four superconducting magnets as shown in Figure 2.8.

The central solenoid is 5.3 m long and 4.5 cm thick and is positioned just outside the inner detector within the calorimeter volume. The central solenoid produces an axial magnetic field of 2T which deflects the paths of charged particles within the inner detector.

There are also three toroid magnets, two end-cap components and one barrel component. The barrel toroid is comprised of eight rounded-rectangular coils spaced evenly around the outside of the calorimeters, which provide a magnetic field of 0.5 T to the muon spectrometer. The innermost edge of the barrel toroid has a diameter of 9.4 m whilst the outermost edge has a diameter of 20.1 m.

Fig. 2.8 Diagram of the ATLAS magnet system showing the windings of the barrel toroid (outermost red tubes) and the end-cap toroid (innermost red tubes) magnets. The central solenoid magnet is the solid red ring located within the calorimeter volume, which is detailed in Section 2.2.2. [35]

The two end-cap toroid magnets also provide a magnetic field of 1 T to the end-cap components of the muon spectrometer. These magnets are also each comprised of eight coils and cover the range $1.6 < |\eta| < 2.7$, whilst the barrel toroid magnets cover the range $|\eta| < 1.4$.

Chapter 3

Level 1 Calorimeter Trigger Noise and Isolation Parameter Optimisation

Collisions at the LHC take place within the ATLAS detector once every ~ 25 ns. Each time the bunches cross, there is a pileup (number of interactions per bunch crossing) of approximately 35 [40] This means there are $\sim 10^9$ interactions every second. Due to the limited bandwidth available to read this data out it is necessary to use triggers in order to only keep events that are likely to contain 'interesting' physics.

The ATLAS trigger system has two levels. The first is a level-1 (L1) trigger which acts at hardware level and uses a subset of the total available information coming only from the calorimeters and the muon system. $\sim 100,000$ events per second pass this first level of trigger. The next level is the high level trigger (HLT) which is a software based trigger that performs a more detailed analysis on the events that have already been refined by the L1 trigger. This refines the dataset to ~ 1000 events per second which are then stored to be used for physics analysis [41].

Events which are not triggered on (i.e. do not pass both aforementioned levels of triggering) are lost forever. Reading out all of the data is already a very difficult task and the instantaneous luminosity that the ATLAS detector is exposed to is expected to increase in the future. Therefore the triggers must be as efficient as possible in order to make sure that as many events as possible that contain interesting physics are stored and readout. This must be achieved in conjunction with keeping the rate of data taking as low as possible in order to operate within the constraints of the limited bandwidth available for data readout. The L1 calorimeter (L1Calo) trigger contributes to the L1 trigger system and so is important at this first stage of reduction in the rate of data-taking. This chapter will cover a study into how different configurations of the L1Calo parameters could possibly be adjusted in order to minimise the rate of data taking whilst keeping the highest possible efficiency of interesting

physics events. These studies are performed using 2016, 2017 and 2018 data in order to look at a variety of data taking conditions and pileup.

It is not possible to redo the online analysis (i.e. actually taking data from the LHC with the ATLAS detector in real time) so offline analyses are performed in order to simulate the variation of online parameters, reconstruct trigger objects (TOBs) and to study whether changes in these lead to an increase or decrease in L1Calo trigger rates and efficiencies. The data set used for efficiency studies is comprised of $Z \rightarrow e^+e^-$ events that are selected using a tag-and-probe method. This method requires using a very tight selection criteria for one electron (the tag electron) such that the efficiency of the identification of the other (probe) electron can be studied. These $Z \rightarrow e^+e^-$ events are chosen because of the relatively large production cross section in conjunction with the easily identifiable signature of the two electrons. To study L1Calo trigger rates, enhanced minimum bias samples are used. These consist mostly of low p_T events which, although not particularly useful for most physics analyses, form the majority of events recorded by ATLAS and so are particularly useful to consider when studying rates.

3.1 L1Calo Reprocessing

In this section the structure of the L1Calo hardware and the software needed to perform offline simulations of the hardware will be summarised. Figure 3.1 shows a summary of the ATLAS trigger and data acquisition system. The L1 trigger systems process information from the muon detector and the calorimeters. Events are stored in a pipeline buffer for ~ 2.5 μs, which is the time it takes in order to make a decision as to whether or not the event will be passed on to the HLT. The HLT uses a large cluster of computers to perform fast-reconstruction algorithms on regions of interest which have been identified at L1. Events passing the HLT are then stored at the CERN computing center for full event reconstruction.

These studies will consider the optimisation of rates and efficiencies of the L1Calo trigger. The aim of this is to study the relationship between these variables and the L1Calo hardware parameters and to see if there is a way to reduce the rate whilst having the smallest possible impact on the efficiency.

3.1.1 The Hardware Processing of L1Calo

A more comprehensive overview of the L1Calo system can be found in reference [43], however a summary is presented here. The ATLAS L1Calo system is a pipelined digital system that is based on custom electronics. It has a fixed latency or dead-time. Approximately

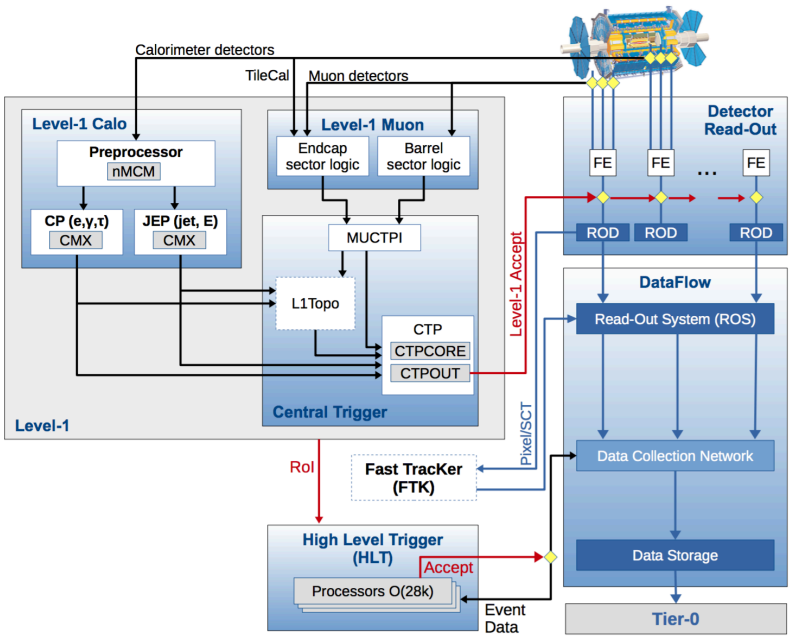


Fig. 3.1 Sumary of the ATLAS trigger and data acquisition system focusing on the trigger. [42]

1.5 μs of dead-time comes from cable transmission delays and a further $\sim$1 μs comes from the actual processing time taken by the central trigger processor (CTP). L1Calo takes input from the electromagnetic calorimeter (ECal) cells and the hadronic calorimeter (HCal) cells. These cells are aggregated to 7200 trigger towers (TTs). Details of the granularity of L1Calo are summarised in Table 3.1. Granularity is measured as a solid angle in $\Delta\eta$ and $\Delta\phi$, where η and ϕ are as described in Section 2.2.

position	$\Delta\eta \times \Delta\phi$
$\lvert\eta\rvert < 2.5$	0.1×0.1
$2.5 < \lvert\eta\rvert < 3.1$	0.2×0.2
$3.1 < \lvert\eta\rvert < 3.2$	0.1×0.2
$3.2 < \lvert\eta\rvert < 4.9$	0.4×0.4125

Table 3.1 Summary of the granularity of L1Calo. η is pseudo-rapidity and ϕ is the azimuthal angle from the beam axis as defined in Section 2.2. These values are taken only in one quadrant ($0 < \phi < \frac{\pi}{2}$).

The three main subsystems of L1Calo are as follows:

- Pre-Processor (PPr): This digitises analogue pulses and then converts the analogue-to-digital counts into an energy. The PPr also associates a TT with the correct bunch crossing (BCID).

- Cluster Processor (CP): Responsible for electron, photon and single hadron identification. It has a granularity of 0.1×0.1 and acts in the region $|\eta| < 2.5$.

- Jet/Energy-sum Processor (JEP): For jet finding and summing the energies of the ECal and the HCal. It has a 0.2×0.2 granularity acting over $|\eta| < 4.9$.

The CPs and JEPs perform operations on the transverse energy (E_T) deposited in ECal TTs, which are as provided by the PPr. Both of these use an overlapping sliding window in order to identify features. This algorithm sums cells within a fixed size rectangular window. The position of this window is then systematically shifted such that the new window is not overlapping with the previous window Looking for local maxima in the total transverse energy contained within the window. More details on this can be found in reference [44]. These results are then sent on to the CTP. The features that the CP algorithms identify are electrons (e), photons (γ) and τ-hadrons. Each CP consists of 4 crates, one for each quadrant of L1Calo, containing 14 cluster processor modules (CPMs). The chips on each CPM are large field-programmable gate arrays (FPGAs). The CPMs carry out the feature identification algorithms and count the number of hits (individual components that detect a signal) in the region covered by each module.

The e/γ algorithm searches for narrow, high-E_T deposits in each ECal TT as provided by the PPr. The characteristics used to separate EM jets from hadronic jets are the transverse isolation of the jet and the amount of 'punch-through' into the HCal. The CP has several sets of isolation and energy threshold conditions, each of which can be chosen independently. Two common merger modules (CMMs), as shown in Figure 3.2, receive information from the back-plane CPMs and jet energy modules (JEMs), which is processed and summed over each crate. The two CMMs use identical hardware but different firmware for the JEMs and the CPMs. The results from the CMMs are then sent to the CTP.

The increased instantaneous luminosity from Run 1 to Run 2 means that there is higher memory occupancy in the L1Calo system. Because of this the data has been condensed into trigger objects (TOBs), which are constructed from individual hits. These contain information on η, ϕ and E_T from the topological trigger. This trigger provides real time measurements of angles between trigger objects and invariant masses. A TOB has a region of interest (RoI) associated with its E_T and also specifies a likelihood of being associated with a particle produced by a collision. Figure 3.3 shows diagrammatically how TTs are summed over all possible 4×4 windows to form TOBs.

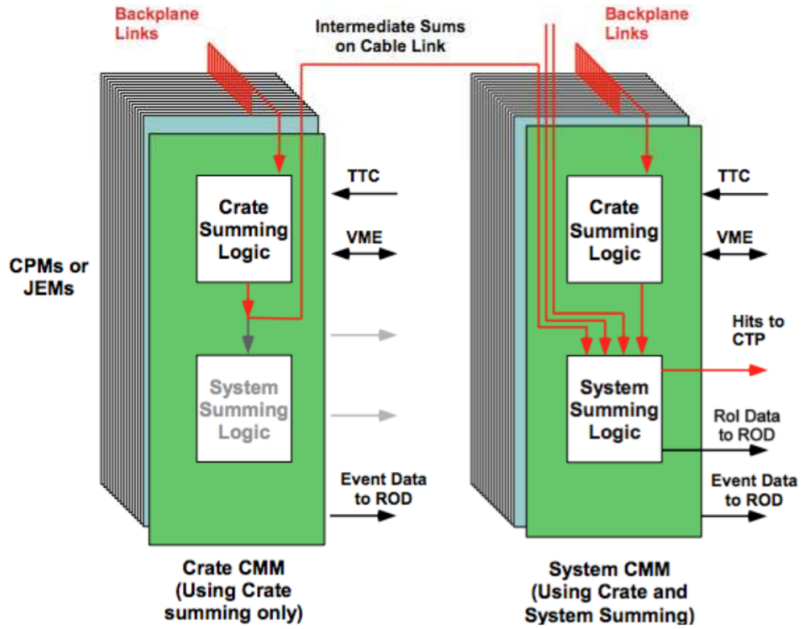


Fig. 3.2 Diagram of the CMMs showing crate-level and system-level functionality.

3.1.2 Simulation of L1Calo

To study the trigger rate and efficiency of L1Calo it is necessary to be able to simulate the TTs and vary the noise and isolation parameters in order to simulate the TOBs. The simulation of the TTs adheres to the following procedure:

- The Condition Objects for LHC computing grid (COOL) database is accessed. The COOL database contains information about the conditions for data taking such as noise settings or dead cell information.
- Retrieve channels which are either dead are disabled.
- Retrieve input TT information in the form of a ‘container’.
- Retrieve ADC digits and copy them into a vector to be reprocessed (i.e. these have already been through ATLAS data processing and this is being performed again to assess the performance with different configurations) with granularity as given in Table 3.1.
- Emulate the finite-impulse-response (FIR), BCID and look up tables (LUTs). This associates a TT pulse with a single bunch crossing.
- Create and fill reprocessed TTs.

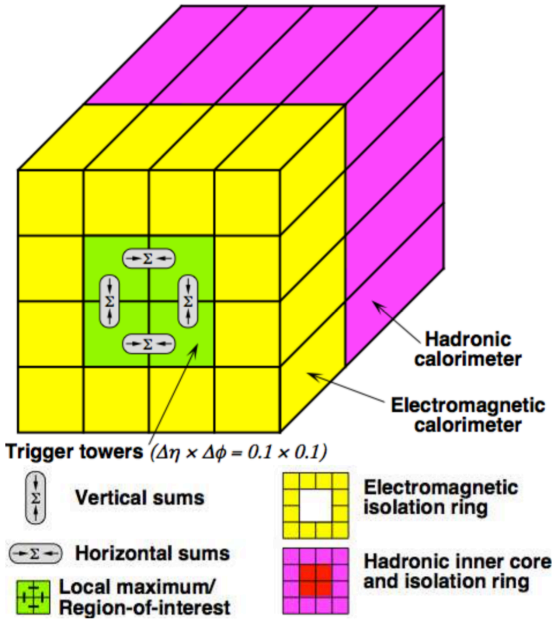


Fig. 3.3 Example of element used for γ, e and τ algorithms. E_T values are summed in each of the possible 2×1 and 1×2 regions in the center of the 4×4 square. In order to find the aforementioned objects, the 2×2 area must have a relatively narrow width and be a local maximum compared to its 8 overlapping nearest neighbours.

TTs have a finite rise time of the order of 50 ns while collisions happen every 25 ns. This means that there can be several bunch crossings over the duration of a single TT pulse. It is therefore very important to be able to associate a pulse to the correct bunch crossing and the method to do this uses a digital pipelined FIR filter as outlined in Figure 3.4.

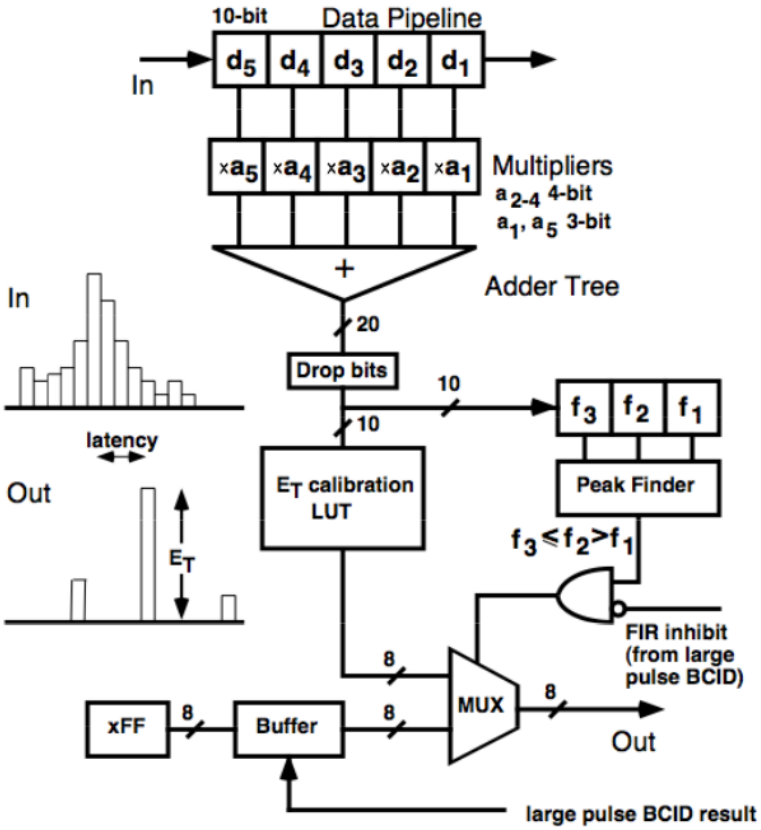


Fig. 3.4 Diagram of the digital pipelined FIR filter showing a typical input and output. [42]

The FIR filter sharpens the TT pulse before using a peak-finder. The peak-finder looks at 5 consecutive samples and multiplies each one by a predefined value, determined by the shape of the pulse with the objective of sharpening the peak, and sums together the resulting values. This sum is then compared with values from the preceding and following bunch crossings to see if there is a maximum. Once a peak is found the LUT, which contains information about the expected amount of background noise, is used to convert the output of the FIR into an energy.

The LUT also provides the calibration for this conversion which requires two values, the slope and the offset, which will be talked about in more detail in Section 3.2.2.

In the LUT simulation we can also alter noise cuts, which are described more in Section 3.2.2, in both the CP and the JEP to study how these effect final trigger rates and efficiencies. Finally, the reprocessed TTs are stored.

Some further steps are required in order to simulate the TOBs from this. TOBs are formed by first iterating over all TTs and summing their transverse energy as well as their angular information in CPM towers. From this the CPMs can then perform algorithms to identify e, γ and τ particles and to count the number of hits in the region of each module.

3.2 Data Selection

Before talking about data selection it is necessary to define some standard ATLAS parlance for data types:

- ESD - Event Summary Data: ESDs contain enough information to simulate the running of parts of the ATLAS experiment. This is raw data that has come from the ATLAS trigger systems, after offline reconstruction has been applied to it.

- DESD - derived ESD: These are filtered ESDs. The filtering depends on the physics requirements of the group using them.

- AOD - Analysis Object Data: AODs are reduced ESDs, with less emphasis on detector related parameters.

- DAOD - Derived AOD: These are filtered AODs. As in the previous case, the filtering depends on the physics requirements of the group using them.

L1Calo specific DAODs are used for 2016 efficiency studies. Because the information stored at AOD level changed between 2016 and 2017, it became necessary to use DESDs for 2017 and 2018 efficiency calculations, which were skimmed for L1 e/γ studies. The rate calculations were performed using enhanced minimum bias ESDs for all years. Events that aren't considered 'interesting' from a physics perspective are filtered out by the triggers. However, if the trigger is to be studied it is important to have more low energy events. Enhanced minimum bias datasets contain this additional information. from A full list of the data sets used is given in Table 3.2 and Table 3.3 as well as the average pile-up for the run ($< \mu >$) the integrated luminosity and the date of the run for completeness.

Dataset Name	$< \mu >$	Luminosity (pb^{-1})	Date
data16_13TeV.00310473.physics_Main.merge .DAOD_L1CALO1.f756_m1708_c1010_m1652	33.3	226.7	13/10/2016
data17_13TeV.00331085.physics_Main.recon .DESDM_EGZ.r10203_r10203	27.4	501.4	30/07/2017
data18_13TeV.00363262.physics_Main.merge .DESDM_EGZ.f997_m1831_f997_m2036	41.8	195.1	09/10/2018

Table 3.2 Summary of data sets used for efficiency calculations. $< \mu >$ is the average pile-up for a run and the date is the day that the data were taken on.

Dataset Name	$< \mu >$	Luminosity (pb^{-1})	Date
data16_13TeV.00302956.physics_EnhancedBias .recon.ESD.r8445_r8540	23.6	337.1	29/06/2016
data17_13TeV.00339070.physics_EnhancedBias .recon.ESD.r10459_r10460	43.3	535.8	26/10/2017
data18_13TeV.00360026.physics_EnhancedBias .recon.ESD.r10897_r10898	35.2	488.4	06/07/2018

Table 3.3 Summary of data sets used for rates calculations. $< \mu >$ is the average pile-up for a run and the date is the day that the data were taken on.

3.2.1 Electron Sample Selection

For efficiency studies a tag-and-probe data sample is used. This is a generic data driven method where a highly abundant dilepton resonance with a very well known distribution is used. A tight selection criteria is applied to one of the electrons (the tag electron) and then the efficiency of the identification of the other electron can be studied. These studies will look at $Z \rightarrow e^+e^-$ events due to the relatively large cross section of this decay. These events are selected from real data.

Selected events must be from luminosity blocks listed on the ATLAS good run list (GRL) and be associated with at least one primary vertex. Luminosity blocks are periods of data-taking assumed to be small enough that the instantaneous luminosity can be considered to be constant. The GRL lists all of the luminosity blocks that are suitable to be used for analysis. Events considered are also required to have a dilepton invariant mass between 66 GeV and 116 GeV in order to be consistent with the mass of the Z boson.

The tag electron is required to fulfill the following criteria:

- Pass one of three separate HLT triggers as listed in Table 3.4. These are selected because they have a very high efficiency (i.e. almost all of the electrons that pass these criteria are real electrons). The trigger names gives the energy threshold that the electron must pass, an identification criterion and an isolation criterion.

- Pass kinematic constraints of $p_{\mathrm{T}}^{e} > 27$ GeV and $|\eta^{e}| < 2.47$ (not including $1.37 < |\eta^{e}| < 1.52$ as this is the transition region between the barrel and end-cap components where there is not full coverage).
- Pass a tight electron identification selection using a loose likelihood (LLH) tool.
- Pass tight isolation requirements.
- Tag electron is matched to a trigger object.

These criteria are summarised in Table 3.4.

Selection	Requirement
HLT	HLT_e26_lhtight_nod0_ivarloose
	HLT_e60_lhmedium_nod0
	HLT_e140_lhloose_nod0
p_{T}	$p_{\mathrm{T}}^{e} > 27$ GeV
$\lvert\eta\rvert$	$\lvert\eta^{e}\rvert < 2.47$
Identification	LooseLikelihood (LLH) Tight
Isolation	Tight
Trigger Matching	Required

Table 3.4 Tag electron selection criteria. The superscript e refers to a quantity possessed by an electron. The HLT names specify the momentum requirement on the electron, the loose likelihood criterion, followed by any isolation criteria.

The probe electron is required to fulfill the following criteria:

- Pass kinematic constraints of $p_{\mathrm{T}}^{e} > 7$ GeV and $|\eta^{e}| < 2.47$ (not including $1.37 < |\eta^{e}| < 1.52$ as this is the crack region, where there is no active detector material).
- Pass a medium selection on the loose likelihood (LLH) tool.
- Pass medium isolation requirements.

These criteria are summarised in Table 3.5. There is no requirement for the object to have passed a trigger as the efficiency of this is what is being studied.

Figures 3.5 to 3.7 show some kinematic distributions for the tag and probe electrons to validate the data sample. Figures 3.5, 3.6 and 3.7 show the reconstructed invariant mass of the dilepton system, the p_{T}^{e} distributions of the tag and probe electrons and the η^{e} distributions for the tag and probe electrons respectively. Figure 3.5 shows a well defined peak centered on the mass of the Z boson, which indicates that this method is properly selecting a pure sample

Selection	Requirement
p_{T}	$p_{\mathrm{T}}^{e} > 7$ GeV
$\lvert\eta\rvert$	$\lvert\eta^{e}\rvert < 2.47$
Identification	LooseLikelihood (LLH) Medium
Isolation	Loose
Trigger Matching	Not Required

Table 3.5 Probe electron selection criteria. The superscript e refers to a quantity possessed by an electron.

of $Z \to e^+e^-$ events. The tag and probe electron distributions show some differences. The tag p_T shows a sharp cut-off where the trigger energy threshold is, whilst the probe electron is less well defined, because its energy can be lower than the trigger threshold.

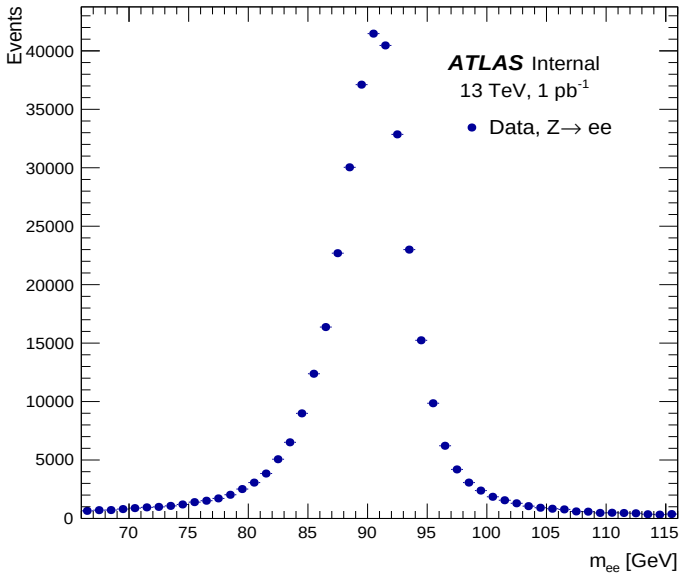


Fig. 3.5 Reconstructed invariant mass of the di-electron system, obtained from the $Z \to e^+e^-$ sample extracted from 2016 data.

Given the reconstruction of a tag electron, the efficiency of reconstructing a probe electron can be measured in the $Z \to e^+e^-$ sample described above. The aim of this study is to measure the trigger efficiency using the offline probe electron matched to an L1 TOB. The TOB is associated to the probe electron by requiring an angular matching. The complete selection on the TOB associated with the probe electron is as follows:

- The angular separation between the TOB and the probe electron is required to be small. It must fulfill:

$$\Delta \mathrm{R} = \sqrt{\Delta\phi^2_{\mathrm{e,TOB}} + \Delta\eta^2_{\mathrm{e,TOB}}} < 0.15 \tag{3.1}$$

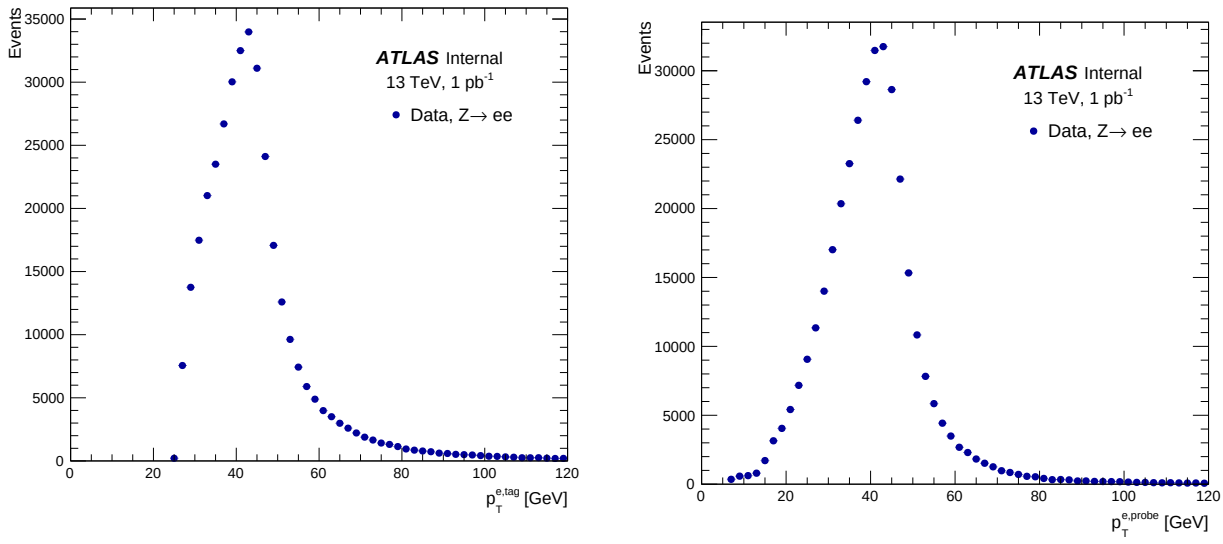


Fig. 3.6 p_{T}^{e} distributions for tag (left) and probe (right) electrons, obtained from the $Z \to e^{+}e^{-}$ sample extracted from 2016 data.

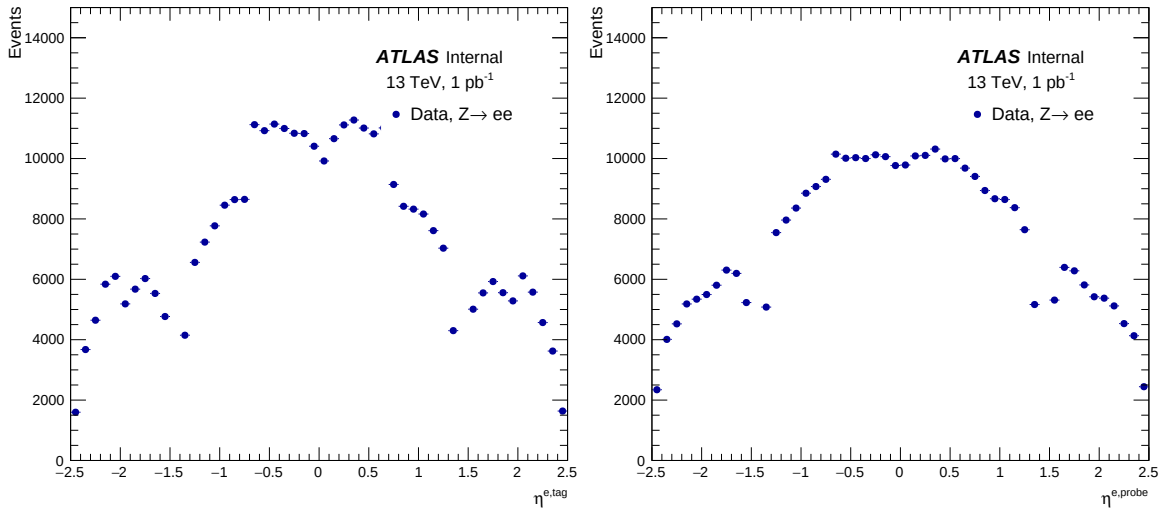


Fig. 3.7 η^{e} distributions for tag (left) and probe (right) electrons, obtained from the $Z \to e^{+}e^{-}$ sample extracted from 2016 data.

where $\Delta\eta_{\mathrm{e,TOB}}$ and $\Delta\phi_{\mathrm{e,TOB}}$ are the differences between the TOB and probe electron in pseudo-rapidity and azimuthal angle respectively.

- The TOB must be reconstructed in a suitable RoI of dimension $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$.

Triggers

The triggers studied cover a range of energy thresholds at the lower end of the viable energy range (where rates and efficiencies are important) and are as follows:

1. EM15HI(2016 data only).
2. EM15VHI(2017 and 2018 data).
3. EM20VHI.
4. EM22VHI.
5. EM24VHI.

Here EM refers to the ECal, the number after the EM tag refers to the minimum transverse energy requirement (in GeV) and the series of letters following that refers to different types and levels of isolation. 'I' refers to the EM ring isolation applied, where the EM energy is within an isolation ring surrounding the core 2×2 trigger towers. 'H' refers to the hadronic core isolation applied (E_{T}-dependent, no isolation for L1 $E_{\mathrm{T}} > 50\,\mathrm{GeV}$). Finally 'V' refers to an η-dependent trigger energy threshold, which is dictated by the variation in η of the energy response (within $+2$ GeV to -3 GeV of the nominal threshold).

3.2.2 Noise and Isolation

Electromagnetic isolation (EM_Isol in Figure 3.8) is defined as the upper limit on the total energy of the 12 0.1 x 0.1 EM towers that surround the 0.2×0.2 EM core region. These regions are as illustrated in Figure 3.3. Figure 3.8 details how the isolation parameters, which can be varied using the reprocessing framework, are defined.

All events falling to the right and below the solid blue line (the shaded region) are accepted. The three isolation parameters that are used and varied in these studies are 'MinCut', slope and offset. MinCut is the minimum isolation energy for which all events, with energy above the trigger threshold, are accepted. Slope and offset then define the isolation cut as a function of the EM cluster energy up until a maximum value which we keep fixed in these studies.

A noise cut of 1000 approximately corresponds to cutting out ~ 1 GeV of noise.

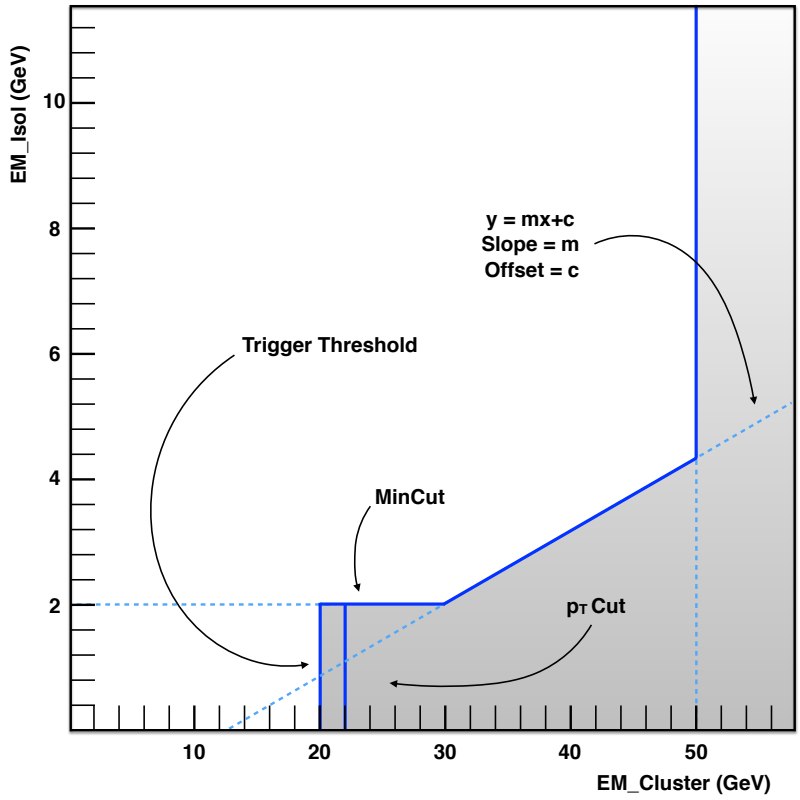


Fig. 3.8 Diagram illustrating how isolation parameters are defined. Everything to the right and below the solid blue line (the shaded region) is accepted

3.2.3 Rate and Efficiency Definitions

The rates being considered here are very simply defined as just the number of times that the trigger in question fires per second. Efficiency must be considered a little more carefully. Efficiency is defined as

$$\varepsilon = \frac{N_{PassTrigger}^{ProbeElectrons}}{N_{Total}^{ProbeElectrons}}. \tag{3.2}$$

In Section 3.3 the efficiency is taken as the average efficiency of reprocessed data (with a custom noise and isolation configuration) over the full range of η with a cut of $p_T >$ trigThresh $+ 4$ GeV. Here 'trigThresh' is the trigger threshold. To illustrate this, Figure 3.9 shows unreprocessed (default) and reprocessed efficiency as a function of p_T for 2017 data. The noise and isolation parameters used are the default ones and the EM24VHI trigger is used (trigger threshold of 24 GeV). For these studies it is desirable to mostly use data which is in the plateau of this distribution so a p_T cut of 4 GeV above the trigger threshold is imposed.

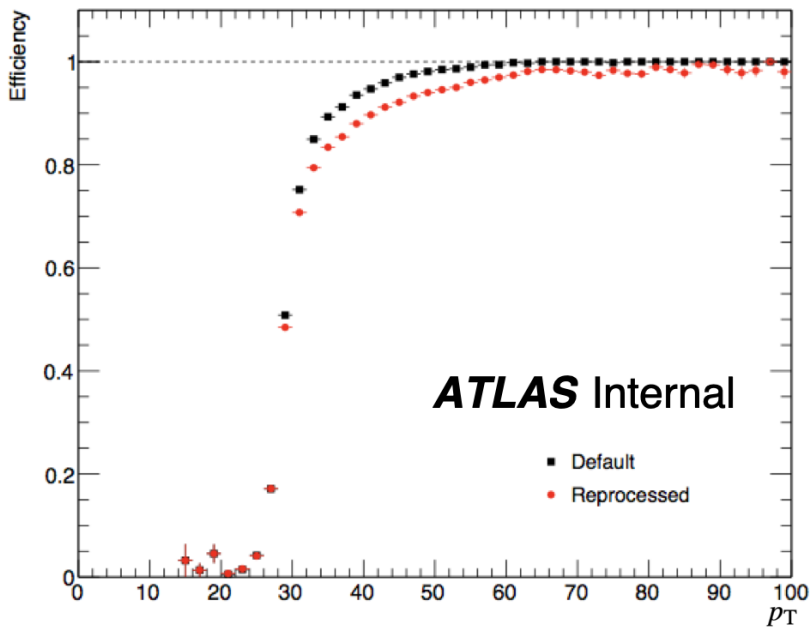


Fig. 3.9 Default vs reprocessed efficiency as a function of p_T for 2017 data using default noise and isolation parameters for the EM24VHI trigger.

Figure 3.10 shows the default vs reprocessed efficiency as a function of η with the same trigger and noise and isolation parameters as in Figure 3.9. The cut $p_T >$ trigThresh $+ 4$GeV has been applied here. The discrepancy between the reprocessed and the unreprocessed can be isolated here is coming from the barrel region ($\eta <\sim 1.3$) only. For the purpose of these studies, a single value is given for the efficiency and the relative difference between these values is of the most importance. Therefore despite this discrepancy, the results in this chapter will still yield good insight on how the various noise and isolation configurations perform relative to each other.

When a single value is required for the efficiency, it is calculated as the average reprocessed efficiency over the full range in η with a cut of $p_T >$ trigThresh $+ 4$ GeV applied to the probe electron.

3.3 Analysis

Results from a study of trigger rates and efficiencies are summarised in this section. The goal of the study is to examine trigger rates and efficiencies for different run conditions (2016, 2017 and 2018) as well as different noise and isolation parameters.

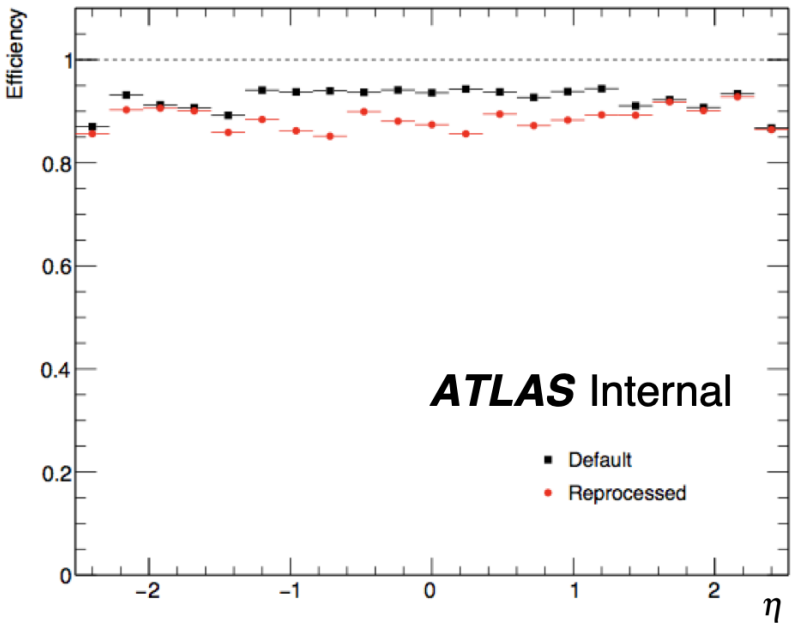


Fig. 3.10 Default vs reprocessed efficiency as a function of η for 2017 data using default noise and isolation parameters for the EM24VHI trigger. A cut in transverse momentum of $p_T >$ trigThresh $+ 4$ GeV is imposed here.

The online noise cut for Run 2 was 4000 and the offset, slope and MinCut were -18, 80 and 20 respectively. A choice was made for these studies to vary the noise parameters between 3000 and 4500 (inclusive) in steps of 500 and to have offset values of -18 or -20 only, slope of 70, 80 or 90 and MinCut of 10 or 20. Initially a winder range of configurations was investigated, however it became obvious that values outside of those shown here were going to have very high rates and/or very low efficiencies so they were discarded.

3.3.1 Results

First, rates and efficiencies are shown for different isolation configurations as a function of noise cut for different triggers. This yields information about the stability of the noise cut (i.e. how much the efficiency fluctuates when the noise cut is varied) as well as illustrating how the noise and isolation parameters effect the rate and efficiency of the trigger.

Results are then shown for three select noise and isolation configurations as a function of trigger energy threshold for different run conditions. Finally rates and efficiencies are presented for different run conditions as a function of trigger energy threshold and then for the same aforementioned configurations.

For 2016 data the EM15VHI trigger was not available and therefore what is being shown here is the EM15HI trigger. The EM15HI trigger does not have any η dependent criteria whilst the EM15VHI trigger does.

Multiple Isolation Configurations as a Function of Noise Cut

The rates shown in Figure 3.12 and in all of the subsequent rates plots in this section have been normalised to the same total area on each plot. This is in order to show the relative difference in rates rather than the absolute value. The relative difference is important as the different samples come from different running conditions and so the absolute value of the rate could be very different.

Figures 3.11 and 3.12 show efficiencies and rates respectively for multiple isolation configurations as a function of noise for 2016 data. For a noise cut between 3500 and 4500 there is little increase in efficiency whilst we see a much larger increase in efficiency between 3000 and 3500. The increase in rates as a function of noise cut is approximately a straight line whilst the efficiency plateaus over the same range. This implies that a noise cut of $\sim$3500 yields a disproportionately higher efficiency relative to the rate for 2016 data.

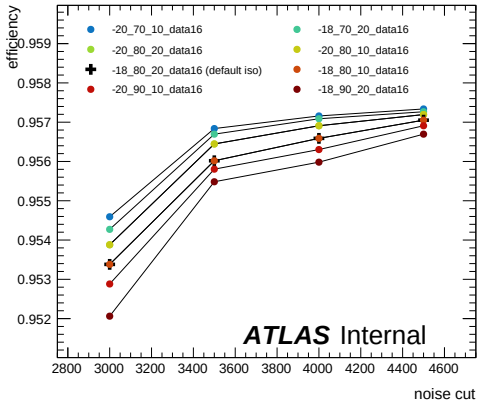
Figures 3.13 and 3.14 show efficiencies and rates respectively for 2017 data. Efficiencies for configurations with a noise cut $\sim$3500 still seem to yield a disproportionately higher efficiency with respect to the rate, which is consistent with the 2016 results.

Figures 3.15 and 3.16 show efficiencies and rates respectively for 2018 data. For the triggers with lower thresholds (EM15VHI and EM20VHI) the plateau between 3500 and 4500 is less pronounced, however it is still prominent for the higher threshold triggers which are more important for 2018 data taking due to the increase in luminosity allowing physics analyses to probe higher energies.

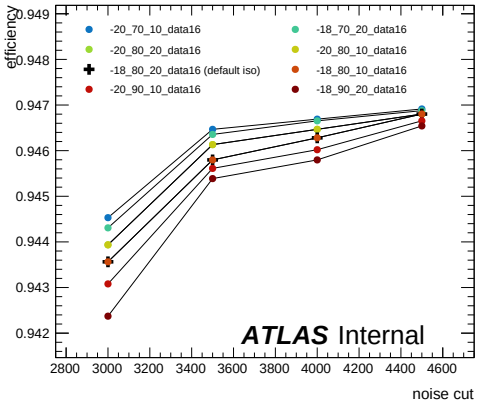
Select Noise and Isolation Configurations as a Function of Trigger

Figures 3.17 and 3.18 show trigger efficiencies and rates as a function of trigger threshold for 2016, 2017 and 2018 data. The three configurations are chosen to be the default (online) configuration and the most and least efficient configurations determined from Figure 3.15d.

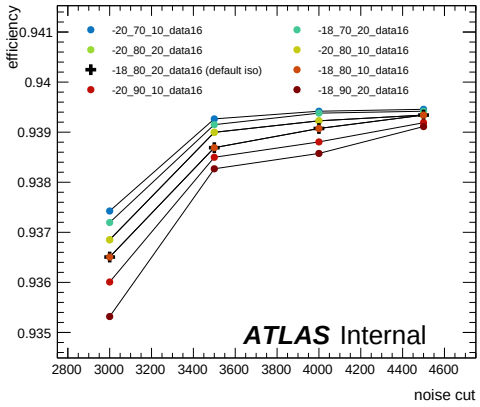
By comparing Figures 3.18b and 3.18c with Figure 3.18a, it can be seen that the rates have proportionally fallen by implementing the η dependent EM15VHI trigger as opposed to the EM15HI trigger used for 2016 data.



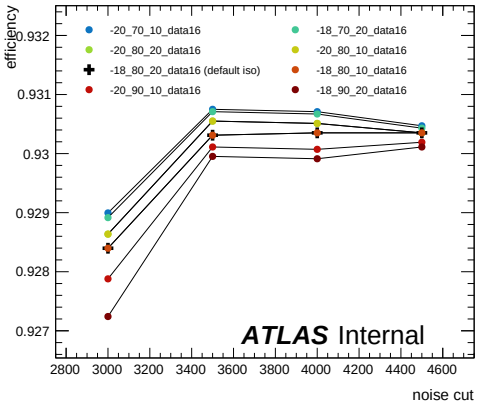
(a) EM15HI



(b) EM20VHI



(c) EM22VHI



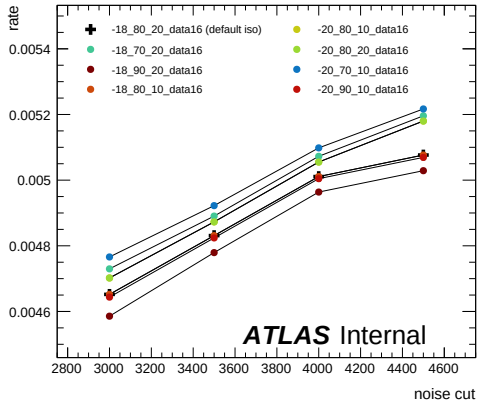
(d) EM24VHI

Fig. 3.11 2016 efficiencies as a function of noise cut for EM15HI (a), EM20VHI (b), EM22VHI (c) and EM24VHI (d). The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.

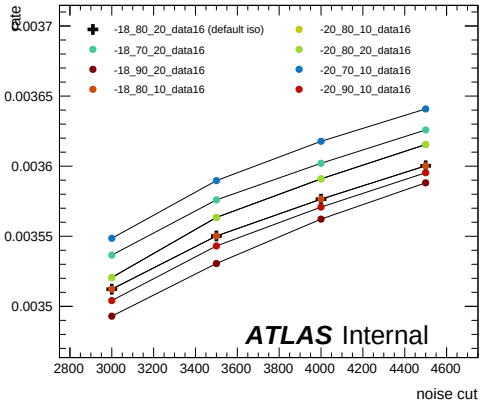
2016, 2017 and 2018 run conditions as a Function of Trigger

Figures 3.19 and 3.20 show trigger efficiencies and rates as a function of trigger threshold. Shown here is a plot for each of the aforementioned selected configurations with contributions from 2016, 2017 and 2018 data superimposed.

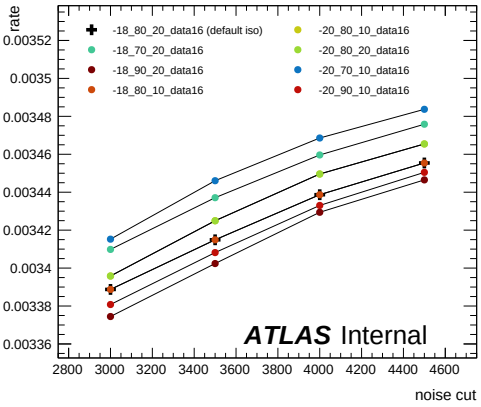
Figure 3.20 illustrates again how the implementation of the EM15VHI trigger has improved trigger rates whilst maintaining efficiency. This is seen in the large spike in rate for 2016 data as this is when the EM15HI trigger was used.



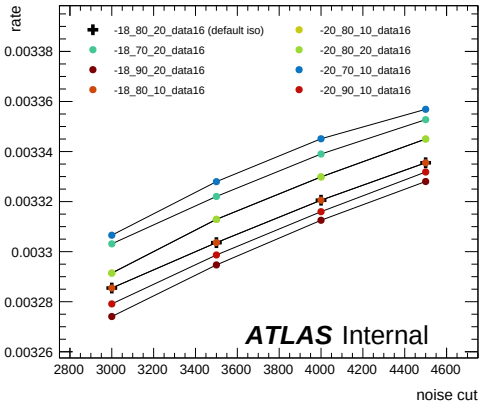
(a) EM15HI



(b) EM20VHI



(c) EM22VHI



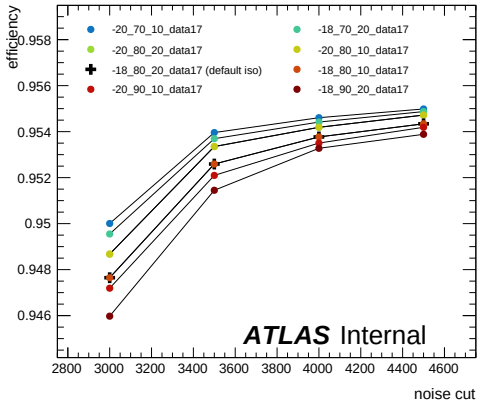
(d) EM24VHI

Fig. 3.12 2016 rates as a function of noise cut for EM15HI (a), EM20VHI (b), EM22VHI (c) and EM24VHI (d). The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.

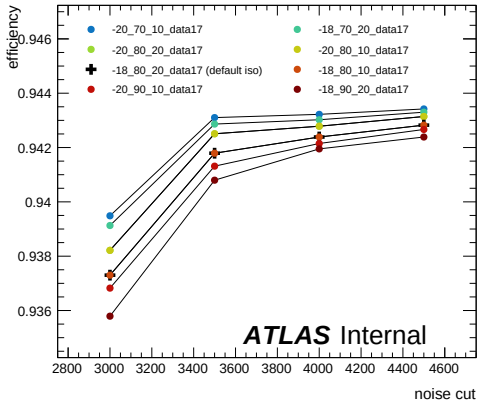
Figures 3.19c and 3.19b, showing results using the default and most efficient noise and isolation configurations, tell us that the EM24VHI yields a higher efficiency for 2018 than 2017 data. The corresponding plots in Figures 3.20c and 3.20b show a proportional decrease in rate.

3.3.2 Summary

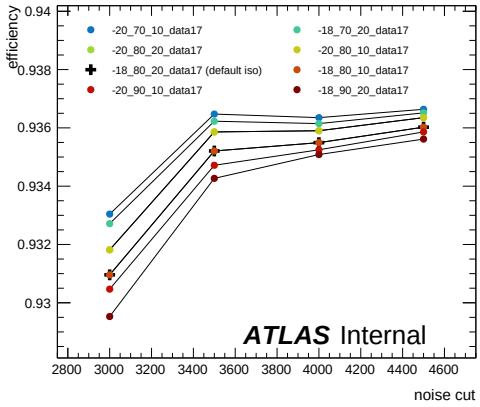
In these studies, the TT rates and efficiencies of the default noise and isolation configuration, as well as the behaviour of several other selected noise and isolation configurations, have



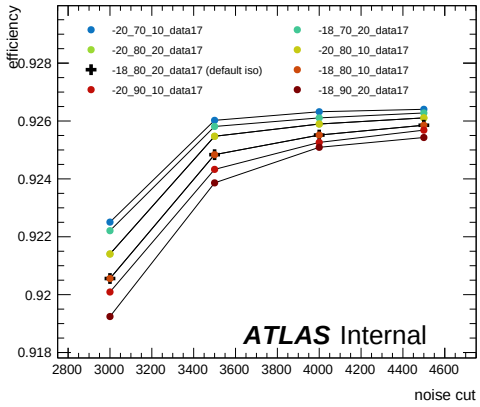
(a) EM15VHI



(b) EM20VHI



(c) EM22VHI



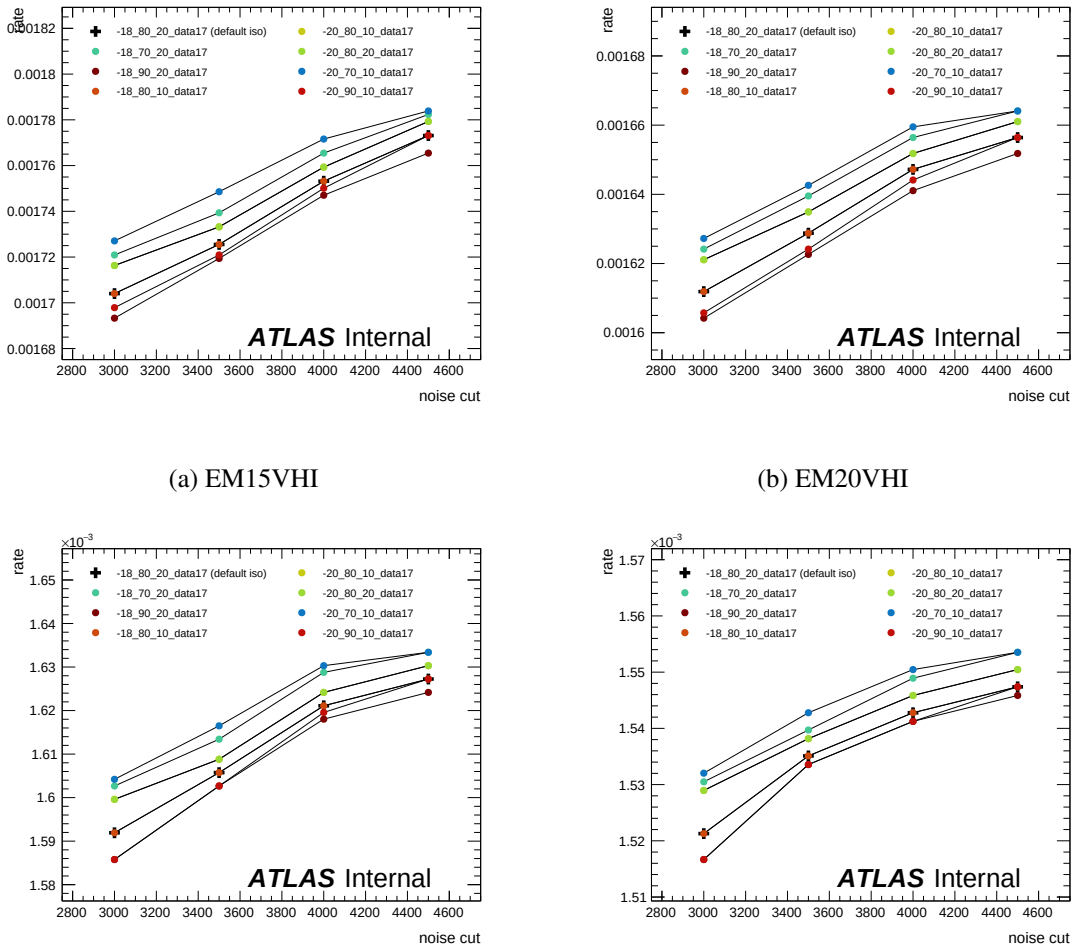
(d) EM24VHI

Fig. 3.13 2017 efficiencies as a function of noise cut for EM15VHI (a), EM20VHI (b), EM22VHI (c) and EM24VHI (d). The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.

been studied. These rates and efficiencies have been presented examining more closely their behaviour as a function of:

- noise cut with eight different isolation configurations superimposed. A plot is shown for each of the four triggers used in this study and for each data taking year in Run 2. This has been optimised by ATLAS and so is expected, but this is confirmed here.

- trigger thresholds with 3 selected noise and isolation configurations superimposed. A plot is shown for each data taking year.



(a) EM15VHI

(b) EM20VHI

(c) EM22VHI

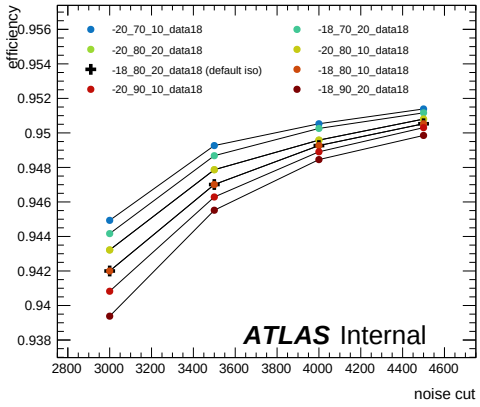
(d) EM24VHI

Fig. 3.14 2017 rates as a function of noise cut for EM15VHI (a), EM20VHI (b), EM22VHI (c) and EM24VHI (d). The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.

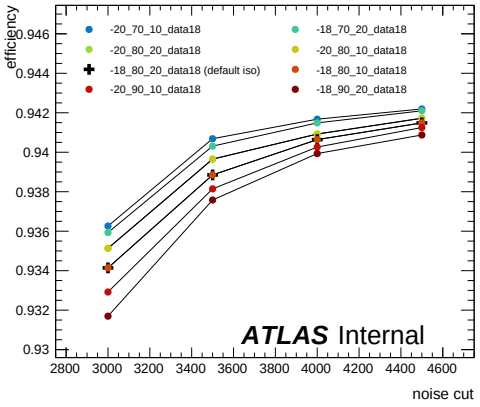
- trigger thresholds with each data taking year in Run 2 superimposed. A plot is shown for each of the three selected noise and isolation configurations.

Key observations include:

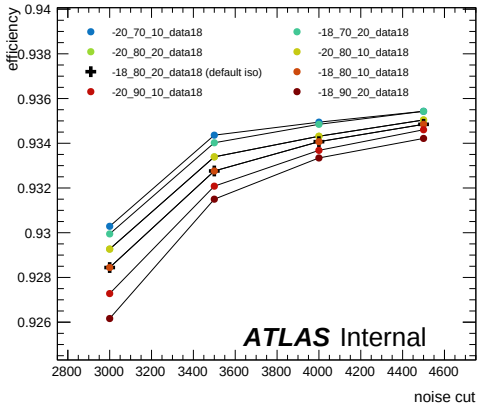
- efficiencies are fairly constant as a function of noise cut for noise cuts between ~ 3500 and ~ 4500. Over this same region, rates increase linearly. This effect is most pronounced for 2016 and 2017 data. This feature becomes less clear but is still present for 2018 data.



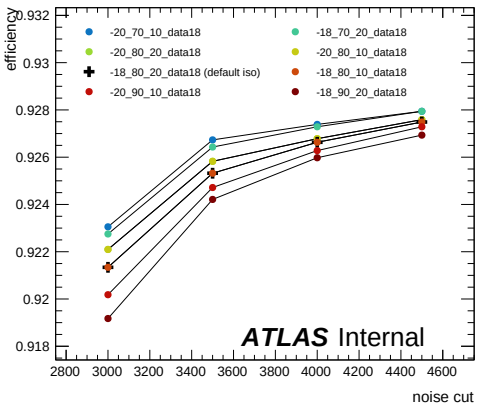
(a) EM15VHI



(b) EM20VHI



(c) EM22VHI

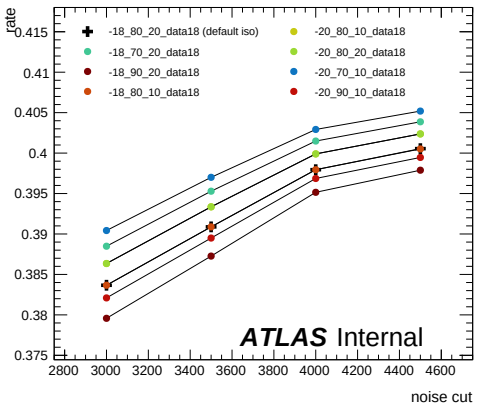


(d) EM24VHI

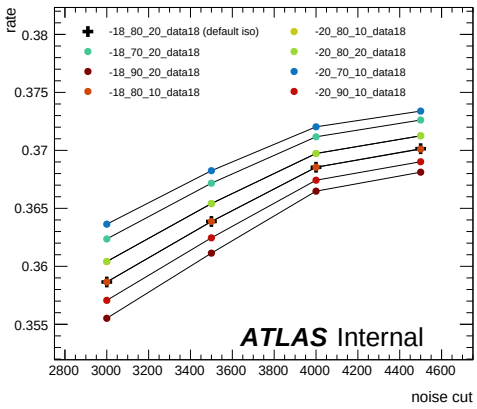
Fig. 3.15 2018 efficiencies as a function of noise cut for EM15VHI (a), EM20VHI (b), EM22VHI (c) and EM24VHI (d). The legend is formatted as $<$ offset $>$ _ $<$ slope $>$ _ $<$ MinCut $>$ _data $<$ runyear $>$.

- the implementation of the η-dependent EM15VHI trigger for 2017 and 2018 data has caused a massive relative decrease in rates compared to EM15HI used for 2016 data with no relative loss in efficiency.

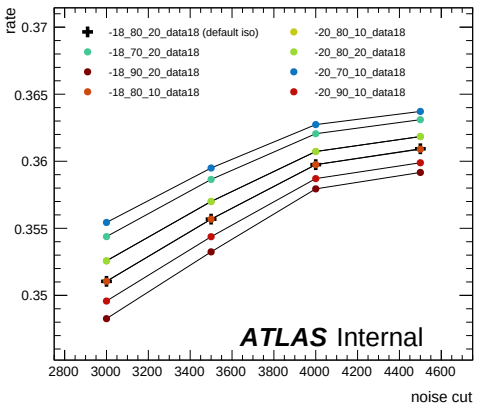
- for the EM24VHI trigger, the overall trigger rate has decreased whilst the overall trigger efficiency has increased for 2018 data when compared to 2017 data. Importantly, this effect is seen for the default noise and isolation configuration, but it is also seen for the most efficient configuration and to a lesser extent in the least efficient configuration.



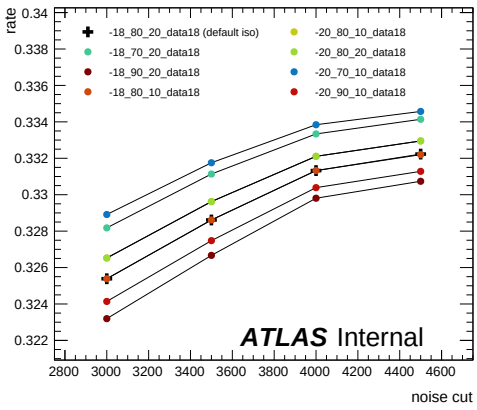
(a) EM15VHI



(b) EM20VHI

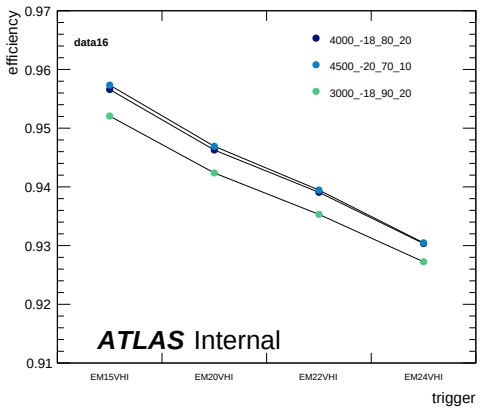


(c) EM22VHI

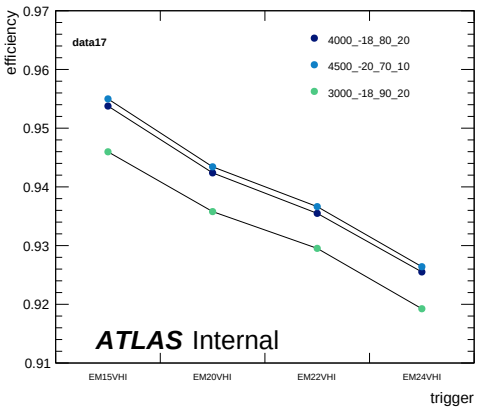


(d) EM24VHI

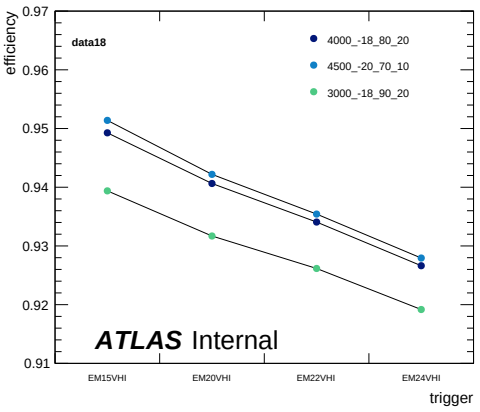
Fig. 3.16 2018 rates as a function of noise cut for EM15VHI (a), EM20VHI (b), EM22VHI (c) and EM24VHI (d). The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.



(a) data16

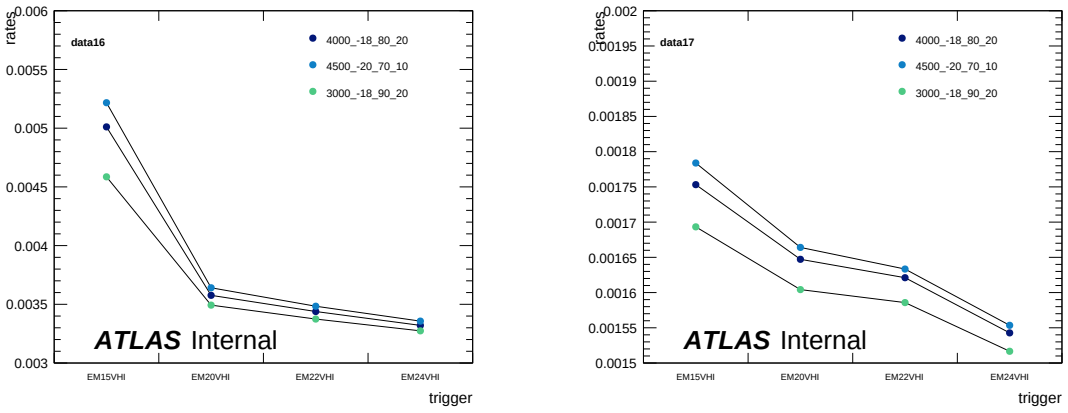


(b) data17



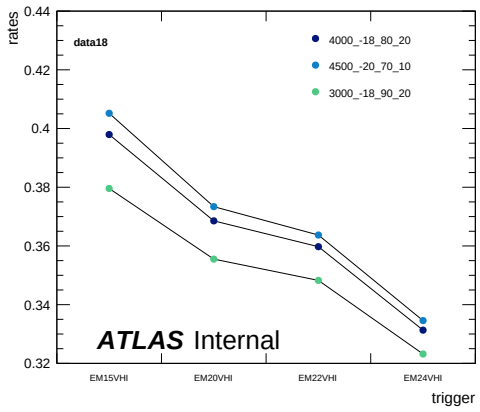
(c) data18

Fig. 3.17 Efficiency as a function of trigger threshold for 2016 (a), 2017 (b) and 2018 (c) data. The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.



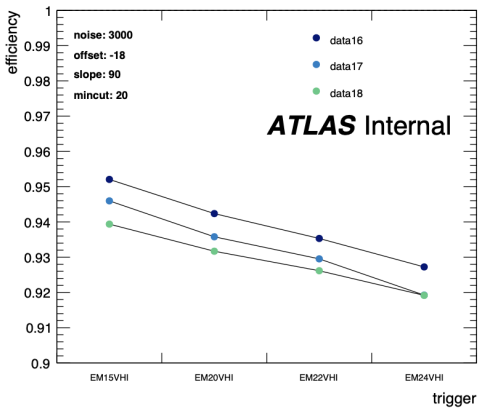
(a) data16

(b) data17

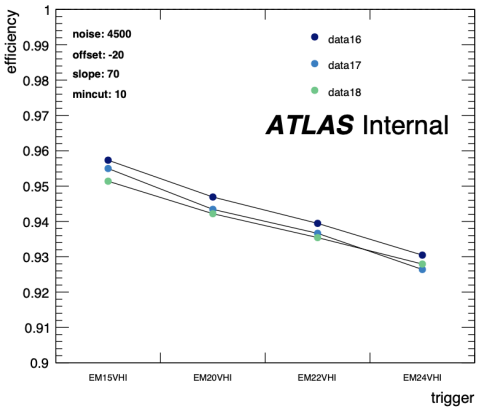


(c) data18

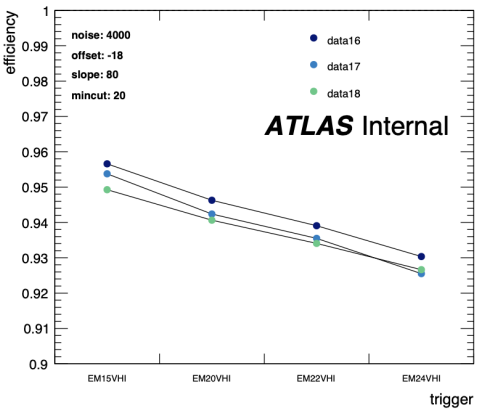
Fig. 3.18 Rate as a function of trigger threshold for 2016 (a), 2017 (b) and 2018 (c) data. The legend is formatted as < offset > _ < slope > _ < MinCut > _data < runyear >.



(a) 3000_-18_90_20

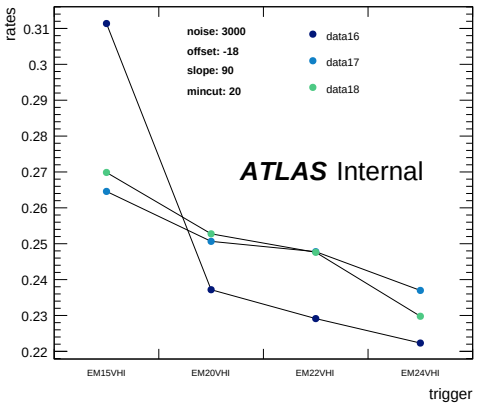


(b) 4500_-20_70_10

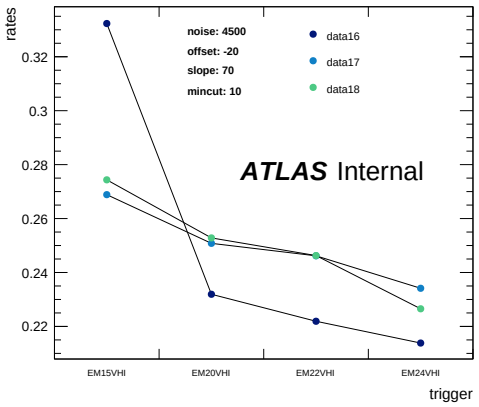


(c) 4000_-18_80_20

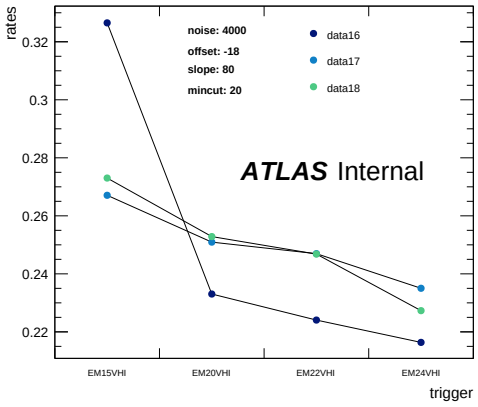
Fig. 3.19 Efficiency as a function of trigger for the least (a) and most (b) efficient configurations and the default (c) configuration.



(a) 3000_-18_90_20



(b) 4500_-20_70_10



(c) 4000_-18_80_20

Fig. 3.20 Rate as a function of trigger for the least (a) and most (b) efficient configurations and the default (c) configuration.

Chapter 4

Event Generation, Experimental Dataset, Object Definitions and Event Selection

4.1 Event Generation

Monte Carlo (MC) simulated data is used in order to estimate the shapes and approximate rate of the various signals and backgrounds in this analysis. There are three MC 'campaigns' which undergo the same physics simulation and are then reconstructed with the pileup profile and run conditions of a given data taking year. mc16a corresponds to the combined 2015 and 2016 data taking years, whilst mc16d and mc16e correspond to the 2017 and 2018 data taking years respectively.

The first step of generating MC events is to simulate the underlying physics making use of MC generators. The propagation of these events through the ATLAS detector is simulated so as to create a model of the actual detector response. Physics objects (e.g. electrons, muons, hadronic jets, etc..) undergo reconstruction based on the detector response in the same way as is done for collision data. In this chapter, the first two steps of modelling the underlying physics and detector simulation will be laid out. The process of object reconstruction will laid out in the subsequent chapter.

4.1.1 Physics Simulation

Parton Distribution Functions

The quarks and gluons which make up a proton can be considered as point like particles and are collectively known as partons, which can then be described by Parton Distribution

Functions (PDFs). An example of a calculation of the cross section of the interaction between two partons is given by

$$\sigma_{a,b\to n} = \sum_{a,b} \int_0^1 dx_a dx_b f_a^{(P1)}(x_a,\mu_F)\, f_b^{(P2)}(x_b,\mu_F)\, d\hat{\sigma}_{ab\to n}(x_a,x_b,\mu_F,\mu_R). \tag{4.1}$$

The $f_i^{(Pi)}(x_i,\mu_F)$ terms correspond to the PDFs, which correspond to a probability distribution of finding a parton, a/b, with a fraction of the total momentum of the proton $x_{a/b}$. This is the short distance cross section calculated from perturbation theory and μ_F and μ_R are the factorisation and renormalisation scales respectively. The factorisation scale can be considered as the choice of energy scale which separates the high energy (perturbative) and low energy (non-perturbative) regimes. This is used to cure UV divergences. The renormalisation scale is used to relate the theory to the energy scale of the experiment.

PDFs depend on the energy scale Q^2, which is the momentum transfer of the interaction. Parton phenomenology is generally dominated by low energy QCD effects and therefore perturbation theory cannot be applied to calculate the PDF. The PDF is therefore determined experimentally using data from other experiments such as HERA [45], the Tevatron [46] and fixed target experiments. However the evolution of these PDFs is still calculated theoretically using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) differential equations [47–49].

Matrix Elements

The production cross-section for a given process can be written as

$$\mathrm{d}\sigma = \sum_{a,b} \frac{1}{x_a x_b s} |M_{ab}|^2 f_a\left(x_a,Q^2\right) f_b\left(x_b,Q^2\right) \mathrm{d}x_a \mathrm{d}x_b \mathrm{d}\Phi, \tag{4.2}$$

where M_{ab} is the matrix element and s is the center of mass energy of the collision. The a and b indices correspond to the initial partons and Φ is the phase space which is then integrated over.

For energies > 1 GeV, the QCD running coupling is low enough that matrix elements can be calculated directly from first principles using perturbation theory. There exist many multi-purpose MC generators which provide SM leading order (LO) matrix elements as well as some BSM extensions. It is also possible to calculate M_{ab} for most SM processes at next-to-leading order (NLO). More recently it has become possible to calculate some next-to-next-to leading order (NNLO) corrections to calculations, but there are used for global corrections rather than in the MC generation itself. It is also possible to calculate

logarithmic corrections without performing the full perturbation theory calculations and so some processes are calculated to leading logarithm (LL), next-to-leading logarithm (NLL) and next-to-next-to-leading logarithm (NNLL).

Parton Showering

As the energy scale decreases, it becomes no longer practical to calculate matrix elements from first principles. In order to model parton showering (PS), a stochastic method is used which employs Markov Chains in order to map the evolution of a shower from the state of the system at a given time. The system probabilistically either emits a quark anti-quark pair or a gluon. The cascade of these forms the shower. In the intermediate energy regime between the exclusively perturbative or exclusively non-perturbative energy scales, the possibility of double counting must be accounted for. Specialised algorithms are used that avoid the problem of double counting. There are detailed in references [50] and [51].

Hadronisation

Hadronisation takes place once the energies are low enough that quarks don't decay any further and colour confinement requires individual them to come together in colour neutral binary or ternary states known as mesons and baryons respectively. This is also a case where explicit calculation of matrix elements from first principles is not possible and so phenomenological methods are required in order to model these processes. There are a wide variety of techniques that can be employed, but covered here are just two amongst the most common.

One common methodology for modelling hadronisation is the Lund String Model [52], where quark anti-quark pairs are considered to be connected by strings. The breaking of these strings allows a mechanism for the formation of additional jets. One of the key predictions of this model is the formation of jets between the two original quarks as well as long the paths of the quarks themselves which is consistent with what is observed in experiments.

Another method is the cluster modelling approach, which relies on the principle of pre-confinement in QCD [53]. Pre-confinement means that quarks and gluons are organised into colour singlets. These colour singlet are then considered to hadronise independently. This can be used to map the distributions of the resulting hadronic jets. Both of these models rely on the tuning of some parameters to experimental data.

Pile-up and the Underlying Event

The underlying event describes any physics which is not a direct result of the hard scattering process and pile-up interactions are those which are not associated with the primary vertex of the hard scattering process. Pile-up at the LHC is modelled using two primary approaches. The first is to carefully simulate the detector to individually model each of the five different types of pile-up; in-time, out-of-time, cavern background, beam halo and beam gas events. In-time pile-up comes from additional pp collisions in the same bunch crossing, whilst out-of-time pile-up comes from pp collisions in different bunch crossings. Cavern background pile-up occurs from background radiation within the detector cavern itself. Beam halo pile-up is caused by the outer parts of the beam interacting with the collimators, which focus the beam as it enters the detector. Finally, beam gas events are caused by the beam interacting with the not-quite-perfect vacuum in the beampipe. Details of these can be found in reference [54]. The second approach makes use of special runs of the LHC in order to model pile-up. These special runs use a random trigger to fully reproduce the run conditions.

The underlying event consists of particles that come from beam-beam remnants and multiple-parton interactions as well as initial and final state radiation. These are modelled using a combination of special runs at the LHC as well as MC modelling.

4.1.2 Detector Simulation

The primary set of tools used for detector simulation at the LHC is the GEANT4 toolkit [55]. There are two main aspects of this simulation. The first is the propagation of the MC generated events through the detector simulation. Then the actual ATLAS running conditions are applied to these events and they can undergo scattering. The response of the detector, which varies depending on the particular component of the detector, is modelled. This is then used for the next stage which is digitisation. Digitisation transforms energy deposits in various components of the detector to actual digital signals.

Fast Simulation

Fast simulation is available within this framework in order to reduce simulation time by approximately an order of magnitude. This is done using the ALTFASTII package [56], which uses a parameterised simulation of the distribution of the calorimeter response to reduce CPU time. As approximately 90% of the particle interactions take place within this part of the detector, this reduces the CPU time for simulation massively and it becomes very important when generating large samples of events.

4.1.3 Dark Matter Signal Models

DM signal models were generated using a fast simulation and a 25 ns bunch crossing configuration. A parametrisation was used for the calorimeter response and GEANT4 was used for all other parts of the detector. These s-channel signal models, as described in Section 1.4, are generated from LO matrix elements with two b-quarks and up to one extra quark or gluon using the MADGRAPHv5-2.6.2 generator [57] interfaced with PYTHIA8.230 [58] for modelling PS, hadronisation and the underlying event. Parton luminosities are taken from the NNPDF30LO PDF set, details of which can be found in reference [59].

Models were produced in a 5-flavour scheme. This means that the hard scale, Q^2, of the process can be much greater than the mass of the bottom quark (m_b), which is important because the signal process involved a $b\bar{b}$ pair. These models were produced at next-to-leading order (NLO) using the DMSimp model [60] in the FeynRules [61] package, which allows for the computation of the Feynman rules of QFT models.

The signal grid (i.e. the choice of mediator mass points) was generated assuming either a scalar (ϕ) or pseudo-scalar (a) mediator decaying into a pair of DM particles of fixed mass $m_\chi = 1$ GeV, which is cosmologically motivated as detailed in Section 1.3.3. Models were generated for mediator masses $m_{\phi/a} = 10,\ 20,\ 50,\ 100,\ 200,\ 300,\ 500$ GeV. This one dimensional signal grid is justified as long as the condition, $m_{\phi/a} > 2m_\chi$ is fulfilled.

4.1.4 Standard Model Backgrounds

SM backgrounds in this analysis are almost all derived entirely from full simulation MC. An overview of the generators, PS and hadronisation methods used are shown in Table 4.1.

Process	MC event generator	PDF	PS and hadronisation	UE tune	Cross-section calculation
V+jets ($V = W/Z$)	SHERPA 2.2.1 [62]	NNPDF3.0NNLO	SHERPA	Default	NNLO [63]
$t\bar{t}+V$	MADGRAPH5_aMC@NLO 2.3.3 [64]	NNPDF3.0	PYTHIA8 [65]	A14 [66]	NLO [67]
tZ	MADGRAPH5_aMC@NLO 2.3.3	NNPDF2.3LO	PYTHIA8	A14	LO
tWZ	MADGRAPH5_aMC@NLO 2.3.3	NNPDF2.3LO	PYTHIA8	A14	NLO [67]
$t\bar{t}$	POWHEG-BOX [68]	NNPDF2.3LO	PYTHIA8	A14	NNLO+NNLL [69–74]
Single top	POWHEG-BOX	NNPDF2.3LO	PYTHIA8	A14	NNLO+NNLL [75–77]
Diboson	SHERPA 2.2.1-2.2.2	NNPDF3.0NNLO	SHERPA	Default	NLO

Table 4.1 Overview of the nominal simulated background samples. The underlying event (UE) tune is a calibration that is based on underlying event or jet radiation observables.

4.2 Experimental Dataset

All of the data for this analysis comes from 2015, 2016, 2017 and 2018 pp collisions recorded by the ATLAS experiment with a center of mass energy $\sqrt{s} = 13$ TeV. Figure 4.1 shows the cumulative integrated luminosity distribution that was delivered by the LHC and also that was recorded by the ATLAS detector. The total combined integrated luminosity used in this analysis is $\mathscr{L} = 139.0 \pm 2.4$ fb^{-1}, which is slightly less than the total amount of recorded data due to the fact that all relevant parts of the detector are required to be fully operational for this analysis. The uncertainty in this value is determined using the method described in reference [78], where a calibration of the luminosity scale is performed using x-y beam separation scans. This uncertainty is determined separately for each data taking year.

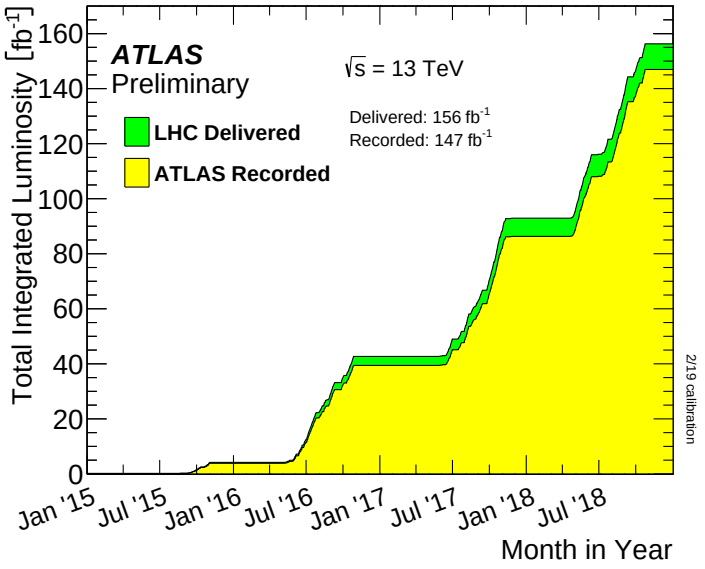


Fig. 4.1 The cumulative luminosity as a function of time recorded by ATLAS (yellow) and delivered by the LHC (green) for all data taking years.

4.3 Object Reconstruction and Definitions

This section will cover the reconstruction of so-called "physics objects". These are electrons, muons, hadronic jets and missing energy. Photons and taus are not of interest in this analysis and so are not discussed here. These objects will be either "baseline" or "signal" objects. Baseline objects are defined first and the signal objects are defined once overlaps between objects (double counting) have been resolved and some additional criteria have been defined.

4.3.1 Jets

Particle Flow (PFlow) jets [79] have become the ATLAS collaboration wide standard for jet reconstruction. These are jets which are reconstructed using a combination of both calorimeter and tracking information. This is as opposed to only using the calorimeter information which used to be the most common way to define jets before Summer 2019. PFlow jets are reconstructed from PFlow objects using the AntiKt4 algorithm [80] with a cone size, $R = 0.4$.

Jets are first calibrated to the EM energy scale [81] and then only jets with $p_T > 20$ GeV and $|\eta| < 2.8$ are selected. These will be called *baseline* jets for the rest of this analysis.

b-tagged Jets

Jets originating from *b* quarks are identified using the DL1r algorithm [82]. This algorithm uses deep learning techniques in order to identify *b*-jets more efficiently. This is done primarily by making use of the characteristic displaced primary vertex and impact paramter of a *b*-jet. The DL1r algorithm takes inputs from the track and calorimeter information which is the fed into a deep neural network which classifies jets.

The DL1r tagger is currently trained on simulated $t\bar{t}$ events as well as exotic Z' events with *b*-jets as the signal and either light flavour jets or *c*-jets as the background. The specific working point (WP) used can then be defined depending on the needs of the specific analysis. The *b*-tagging WP chosen for this analysis is such that 77% of b-tagged jets are correctly identified in MC. Details of the rejection efficiency of non *b*-tagged jets can be found in reference [83].

4.3.2 Leptons

A lepton, in this analysis, refers to either an electron or a muon or their corresponding anti-particles. In general, the treatment of taus is very different as they will leave a signal in the hadronic calorimeter and can be difficult to separate from hadronic jets.

Reconstruction of electrons relies on the clustering of energy deposits in the ECal. These deposits must also be associated with a track in the ID. The energy deposits are then calibrated to the EM scale. The baseline requirement for electrons in the analysis is $p_T > 7$ GeV, $|\eta| < 2.47$ and they must pass the `LooseAndBLayerLLH` requirement [84]. This is a log-likelihood based requirement that also includes information from the IBL.

Muons are also reconstructed requiring a track in the ID and this track must match with another track in the MS. Baseline muons are defined to have $p_T > 6$ GeV, $|\eta| < 2.7$ and

must also pass the Medium identification criteria [85]. These criteria have been designed to minimise the systematic uncertainties associated with muon reconstruction and calibration.

4.3.3 Missing Transverse Energy

Missing transverse energy (E_T^{miss}) in this analysis is calculated as the absolute magnitude of the negative vector sum of all energy deposits and tracks in the various subsystems of the detector. It is calculated after matching between tracks and energy deposits is done in order to minimise the possibility of double counting. This vector sum can be written as

$$\mathbf{E}_{\mathrm{T}}^{\text{miss}} = \mathbf{p}_{\mathrm{T}}^{\text{miss},e} + \mathbf{p}_{\mathrm{T}}^{\text{miss},\gamma} + \mathbf{p}_{\mathrm{T}}^{\text{miss,jets}} + \mathbf{p}_{\mathrm{T}}^{\text{miss},\mu} + \mathbf{p}_{\mathrm{T}}^{\text{miss,soft}}, \tag{4.3}$$

where $\mathbf{p}_{\mathrm{T}}^{\text{miss},e} = -\mathbf{p}_{\mathrm{T}}^{e}$ and so forth. E_T^{miss} is the absolute magnitude of this quantity. This analysis uses a tight E_T^{miss} WP to ensure that the E_T^{miss} actually comes from a particle and does not arise due to mismeasurement. This WP requires that the leading jet (the jet with the largest transverse momentum in the event) p_T must be greater than 30 GeV.

An important background in this analysis is $Z(\rightarrow \nu\nu) + bb$ and, in order to estimate the size of the contribution of this process in the signal regions, a two lepton control region is used. the two lepton system is used as a proxy for the $\nu\nu$ system, where the two leptons are added to the E_T^{miss} calculation in Equation 4.3. From this point onward, it will be assumed that the E_T^{miss} calculation includes the p_T of the two lepton system wherever two leptons are considered. A new variable, $\tilde{E}_T^{\text{miss}}$, is used for instances where two leptons are considered but the p_T of the di-lepton is not added to the E_T^{miss} calculation.

4.3.4 Overlap Removal

Overlap removal (OR) algorithms are techniques applied to reconstructed objects that ensures that the same particle is not reconstructed as two different objects for the analysis. This is a set of algorithms that determine which object should take precedent in cases where more than one set of object criteria is passed. Details of the OR techniques applied in this analysis can be found in reference [86]. To maximise the possible yield of b-jets in this analysis, an OR method has been used which favours the choice of a b-jet for jets with $p_T < 100$ GeV.

4.3.5 Signal Objects

In addition to the baseline requirements, some objects are also required to pass additional criteria which will define them as *signal* objects for the rest of this analysis.

Jets with $20 < p_T$ [GeV] < 60, $|\eta| < 2.4$ and are required to have a jet vertex tagger (JVT) score of $w_{\text{JVT}} > 0.5$. This is to reject jets which have originated from pile-up interactions. This metric is defined as

$$w_{\text{JVT}} = \frac{\sum_k \Delta R(\text{jet}, k) p_{\text{T}}^k}{\sum_k p_{\text{T}}^k}, \tag{4.4}$$

where the index, k, runs over all of the constituents of the jet [87] and ΔR is as defined in Section 2.2. This metric therefore corresponds to the width of the jet. These requirements collectively correspond to the TightJVT WP for PFlow jets.

Signal electrons must then pass some additional isolation and identification requirements:

- $|z_0 \sin\theta| < 0.5$
- $\left|\frac{d_0}{\sigma(d_0)}\right| < 5$

Here, z_0 and d_0 are the longitudinal and transverse (with respect to the beam axis) impact parameters respectively and $\sigma(d_0)$ is the uncertainty in d_0. In order to select electrons that are from the hard scattering process, a requirement of $p_T > 20$ GeV is also imposed. Lastly, in the case where there are two electrons in the event, the leading electron is required to be matched with the electron that fired the trigger and is required to have $p_T > 27$ GeV in order to ensure that the trigger is fully efficient.

Signal muons require no further identification or isolation requirements to their baseline counterparts. Similarly to signal electrons, signal muons are required to pass the following requirements on their impact parameters:

- $|z_0 \sin\theta| < 0.5$
- $\left|\frac{d_0}{\sigma(d_0)}\right| < 3$

The p_T requirement on signal muons depends on the trigger which was fired for the event. The ATLAS E_T^{miss} triggers do not take into account the muon spectrometers and so high p_T muons can cause a E_T^{miss} trigger to be fired. When a E_T^{miss} trigger is fired, the muon is required to have $p_T > 20$ GeV in order to reject fake leptons. Fake leptons could arise from, for example, a jet being misidentified as an electron. In the case that a lepton trigger is fired, the requirement is $p_T > 27$ GeV in order to ensure that the trigger is fully efficient.

4.4 Event Selection

In this section the analysis preselections will be presented as well as details of the trigger strategy used.

4.4.1 Initial Selection

To keep the size of the data sets small enough to be able to process quickly, skimmed derivations of the full data sets are used. The derivation used for this analysis requires events to pass one of the following criteria:

- The scalar sum of the p_T of all of the jets in the event, or the total hadronic energy, $(HT) > 150$ GeV
- One lepton or photon with $p_T > 100$ GeV
- Two leptons or two photons with $p_T > 20$ GeV

These are collectively known as the 'SUSY1' filtering criteria. For the remainder of this thesis, these requirements are assumed to have been applied unless explicitly stated otherwise.

4.4.2 Event Cleaning

This following is a list of criteria applied to the data and the MC in order to remove poorly measured and/or pathological events:

- Data Event Cleaning: A veto is placed on luminosity blocks which have some relevant sub-detector functioning poorly. Luminosity blocks are periods of data-taking assumed to be small enough that the instantaneous luminosity can be considered to be constant.
- Vertex Selection: Events are required to have a primary vertex. This means that it should be possible to reconstruct the origin of the event.
- Jet Cleaning: This is a process designed to mitigate the acceptance of events where jets arise from sources other than the collision. For example, jets can arise from cosmic muons or fake signals in the calorimeter coming either from noise bursts or from the presence of coherent noise. The set of cuts applied has an almost 100% efficiency of selecting signal jets and rejecting non-collision jets.
- Bad Muon Veto: Events that satisfy $\frac{\sigma\left(\frac{q}{p}\right)}{\frac{q}{p}} > 0.2$ before overlap removal are rejected, where q is the charge of the muon, p is the momentum and $\sigma\left(\frac{q}{p}\right)$ is the uncertainty in the ratio of these two quantities.

4.4.3 Trigger Strategy

This analysis is targeting events that are expected to have zero leptons and a large amount of missing energy and so a preselection signal requirement is made that events must pass HLT-E_T^{miss} triggers. The HLT-E_T^{miss} trigger used is dependent on the data taking period. The details of these triggers can be found in Table 4.2.

Data Period	HLT_xe Trigger
2015	HLT_xe70_mht
2016 A-D3	HLT_xe90_mht_L1XE50
2016 D4-F1	HLT_xe100_mht_L1XE50
2016 F2-	HLT_xe110_mht_L1XE50
2017 B1-D5	HLT_xe110_pufit_L1XE55
2017 D6-	HLT_xe110_pufit_L1XE50
2018 B-C5	HLT_xe110_pufit_xe70_L1XE50
2018 C6-	HLT_xe110_pufit_xe65_L1XE50

Table 4.2 A list of the lowest unprescaled E_T^{miss} triggers used in zero lepton regions which triggered on by E_T^{miss}. The names of the triggers are formatted to indicate what level of trigger is considered (HLT for all here), the energy threshold of the trigger (in GeV), what identification requirement must be passed and the requirement on the missing energy threshold of the L1 trigger that initially fired (in GeV).

Some regions of phase space used in this analysis require leptons, so lepton triggers are used in these regions. These are as listed in Table 4.3

Data Period	Electron Trigger	Muon Trigger
Data 2015	HLT_e24_lhmedium_L1EM20VH	HLT_mu20_iloose_L1MU15
Data 2016-18	HLT_e26_lhtight_nod0_ivarloose	HLT_mu26_ivarmedium

Table 4.3 A list of single-lepton triggers used. The lowest unprescaled trigger used during each data-taking period is shown. The names of the triggers are formatted to indicate what level of trigger is being considered (HLT for all), the energy threshold (in GeV) of the object (either an electron, e, or a muon, mu) as well as any identification and isolation requirements. The 2015 triggers also required a L1 trigger to be fired.

The so-called 'turn-on curves' for the lowest unprescaled triggers are shown in Figure 4.2 and Figure 4.3. The turn-on curve shows the transition of the trigger from hardly firing at all to being 100% efficient, as a function of a kinematic variable. The reference trigger used here is a single muon trigger with a requirement that the p_T of the muon be > 30 GeV so that the trigger is fully efficient. Because the muon trigger is known to be fully efficient under

this requirement, the E_T^{miss} trigger turn-on curve is divided by the muon trigger turn-on curve. Doing this allows us to see at what point the E_T^{miss} turn on curve becomes fully efficient also. This trigger is used as the online E_T^{miss} calculations are made only using calorimeter information, so muons would be equivalent to missing energy in these calculations. The x-axis is labelled as $E_T^{\text{miss(muons inv.)}}$ to reflect this. The only other cuts applied here are those made at derivation level which are summarised in Section 4.4.1.

It is clear from this simple one dimensional study, shown in Figure 4.2 and Figure 4.3, that in order to be safely on the plateau of the trigger turn-on curve selection of $E_T^{\text{miss}} > \sim 220$ GeV would be required. However, for some of the low mass mediator signal models, it is anticipated that being able to explore a lower E_T^{miss} phase-space could be important. For this reason a study is shown here that shows how it is possible to go lower in E_T^{miss} if the correlation between E_T^{miss} and p_T of the leading jet ($p_T(j_1)$) is utilised for a 2D cut on a function of these two variables.

Figure 4.4 and Figure 4.5 show the 2D trigger efficiency when considering both E_T^{miss} and $p_T(j_1)$. From this it can be seen that there is a large region of phase-space where the E_T^{miss} is fully efficient for lower values of E_T^{miss} provided that a requirement is simultaneously placed on $p_T(j_1)$.

It was found that the following requirement on E_T^{miss} guarantees plateau in the trigger efficiency curves as a function of $p_T(j_1)$:

$$E_{\text{T}}^{\text{miss}} > \frac{5000}{p_{\text{T}}(j_1) - 20} + 160,$$

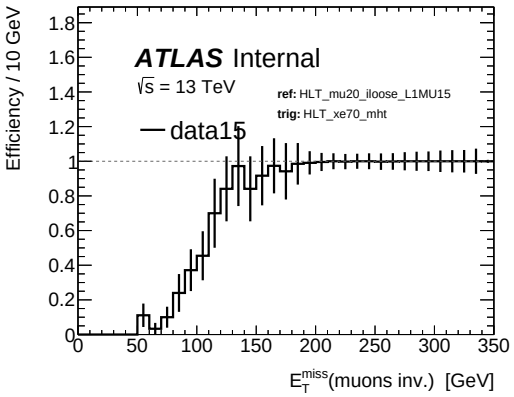
where the units of all the numbers and variables are in GeV apart from 5000, which is in GeV^2. This can be rearranged to

$$\frac{1}{5000}(p_{\text{T}}(j_1) - 20)(E_{\text{T}}^{\text{miss}} - 160) > 1.$$

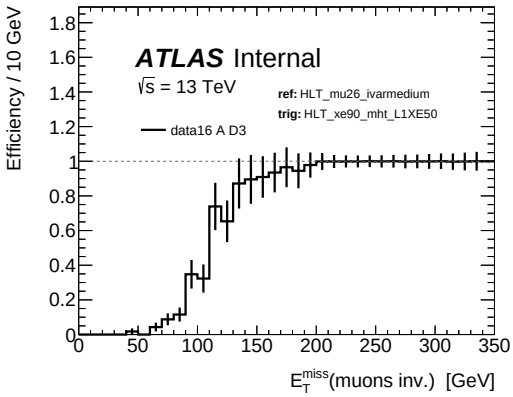
The left hand side of this inequality is defined as trigPlateau.

The line representing this inequality has been drawn onto each of the 2D E_T^{miss} vs $p_{\text{T}}(j_1)$ plots for the relevant lowest unprescaled trigger in a given period. Everything above and to the right of this line will fall into the plateau of the trigger efficiency turn-on curves.

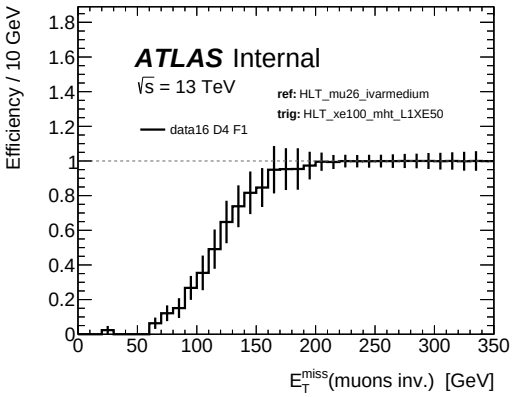
Figure 4.6 and Figure 4.7 show the trigger efficiency as a function of E_T^{miss} with a selection of trigPlateau > 1 applied. As well as this, selections of $\frac{p_T(j_1)}{HT} > 0.7$ and object based E_T^{miss} significance > 7 (defined in Section 4.4.4 and in reference [88]) are also applied due to their correlation with E_T^{miss}. These selections are applied at preselection level for all regions in this analysis. Apart from a few bins with very low statistics, all bins are fully efficient after this selection is applied.



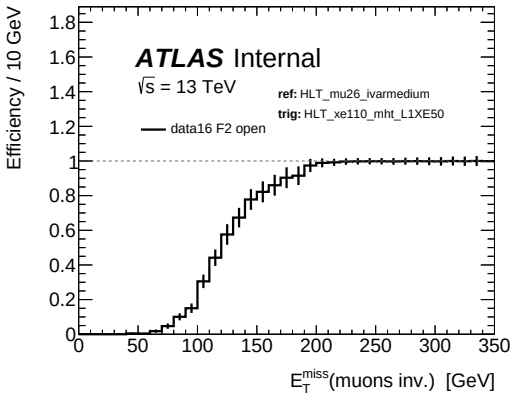
(a) Data 2015



(b) Data 2016 A-D3



(c) Data 2016 D4-F1



(d) Data 2016 F2-

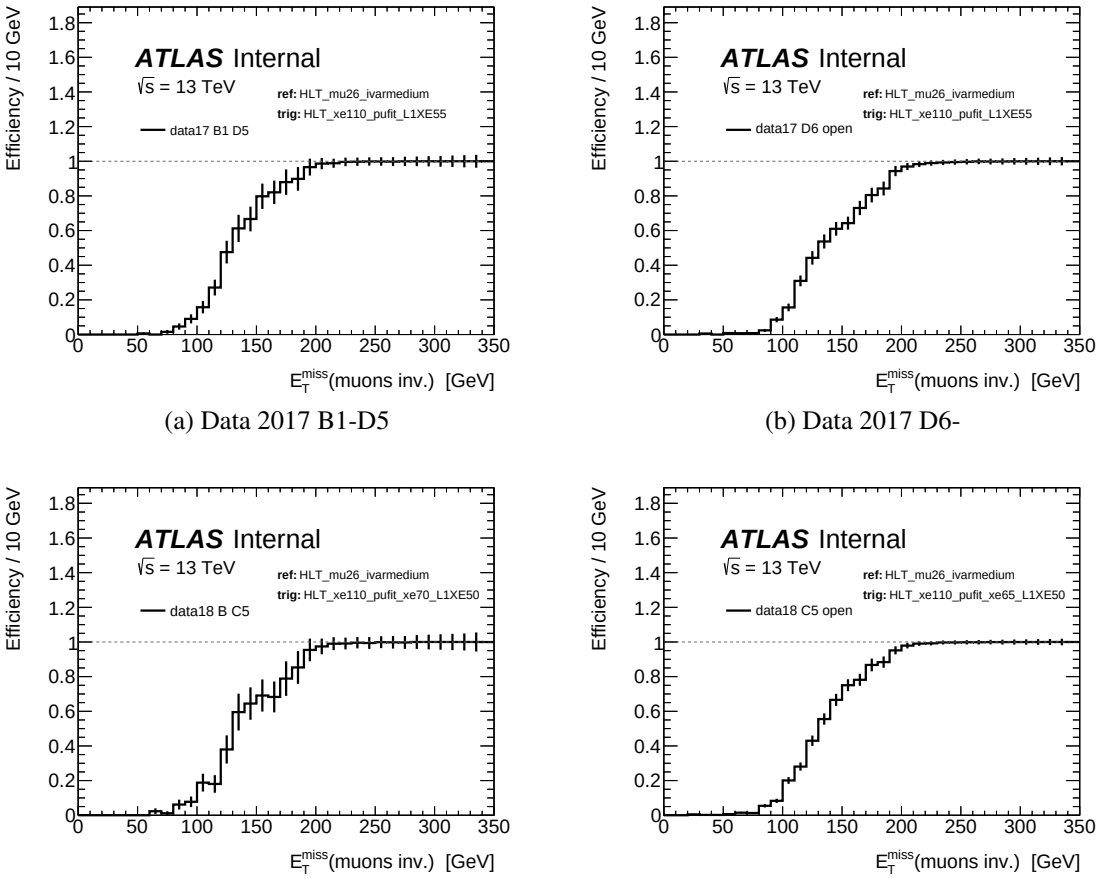
Fig. 4.2 Lowest unprescaled E_T^{miss} trigger turn-on curves for each data taking period from 2015 to the end of 2016. Uncertainties are statistical only.

4.4.4 Discriminating Variables

Part of the challenge of this analysis is to be able to discriminate the signal models from background SM processes which leave very similar signatures in the ATLAS detector. For this reason it is necessary to define a set of discriminating variables, which can be used to take advantage of any small differences that do exist, to achieve the greatest possible sensitivity to the DM signal models. This section details all of the kinematic variables used in this analysis.

Mass Scale Variables

Mass scale variables connect observables to the mass scale of DM signals. The following is a list of mass scale variables used in this analysis:



(a) Data 2017 B1-D5

(b) Data 2017 D6-

(c) Data 2018 B-C5

(d) Data 2018 C6-

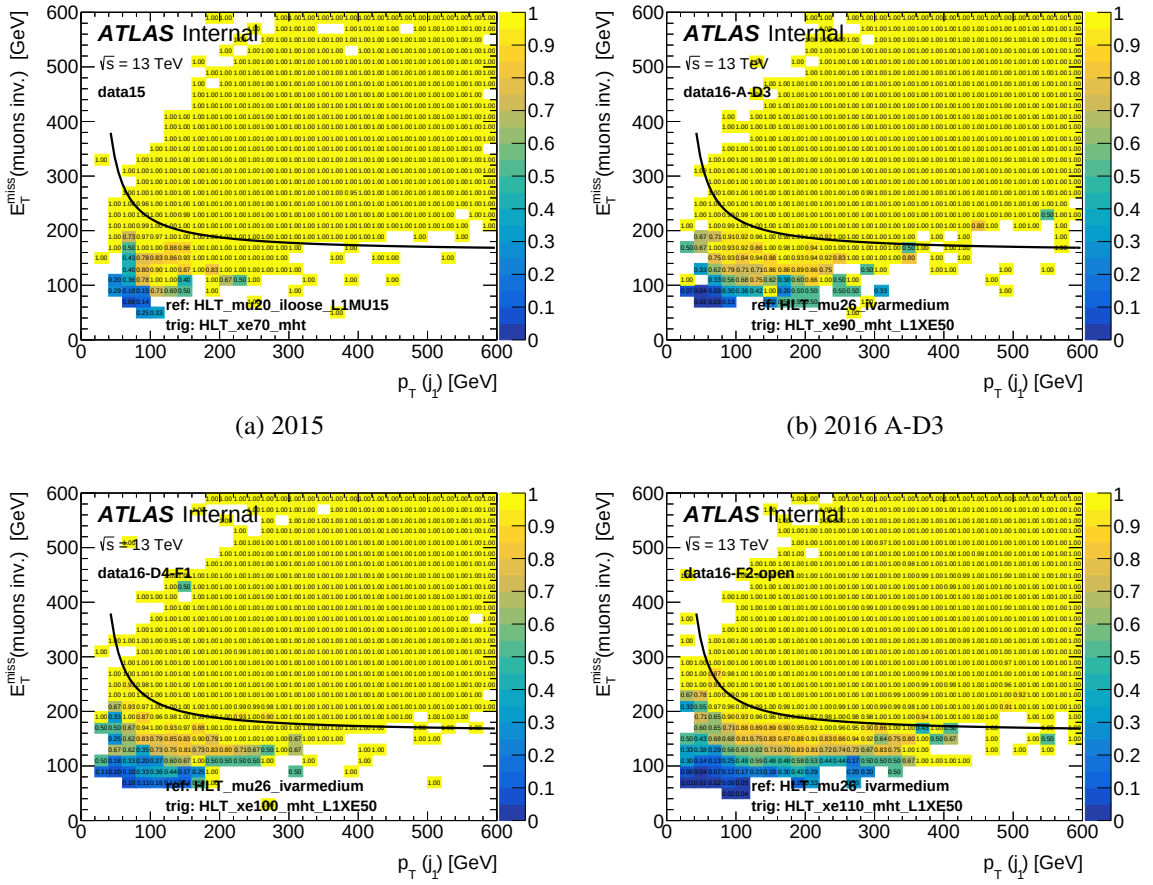
Fig. 4.3 Lowest unprescaled E_T^{miss} trigger turn-on curves for each data taking period from 2017 to the end of 2018. Uncertainties are statistical only.

- E_T^{miss}: Defined in Section 4.3.3.

- Object Based E_T^{miss} Significance (S): As described in Reference [88], object based E_T^{miss} significance is defined as

$$S = \frac{\left|\mathbf{E}_{\text{T}}^{\text{miss}}\right|}{\sqrt{\sigma_{\text{L}}^2\left(1-\rho_{\text{LT}}^2\right)}}, \tag{4.5}$$

  where σ_{L} is the estimated total longitudinal momentum resolution of all jets and leptons at a given η and a given p_T. ρ_{LT} is the correlation factor between each objects longitudinal and transverse momentum resolution. This variable is used to discriminate between real E_T^{miss} (i.e. E_T^{miss} that comes from invisible objects like neutrinos or BSM



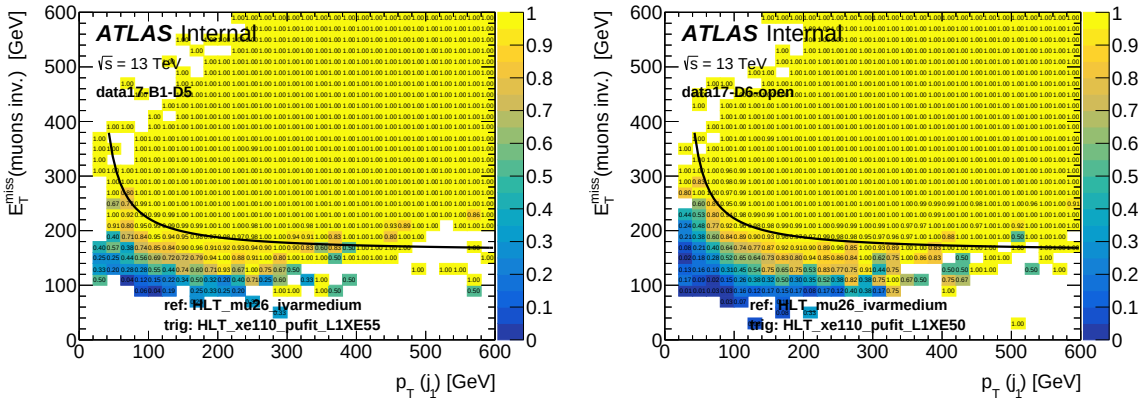
(a) 2015

(b) 2016 A-D3

(c) 2016 D4-F1

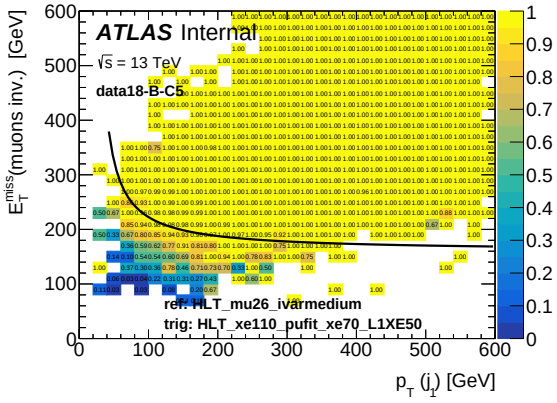
(d) 2016 F2-

Fig. 4.4 2D trigger efficiency E_T^{miss} vs $p_{\text{T}}(j_1)$ for each data taking period from 2015 to the end of 2016. Superimposed onto each plot is the proposed 2D selection that would ensure the trigger is fully efficient.

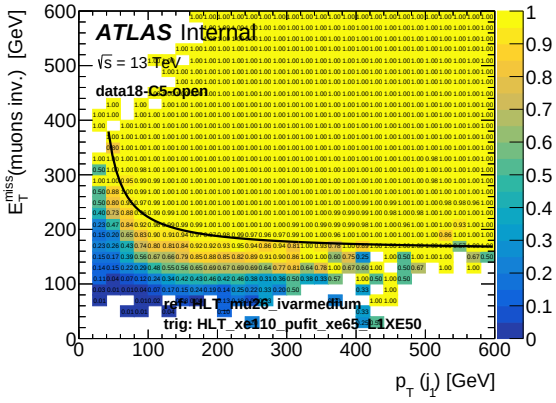


(a) 2017 B1-D5

(b) 2017 D6-

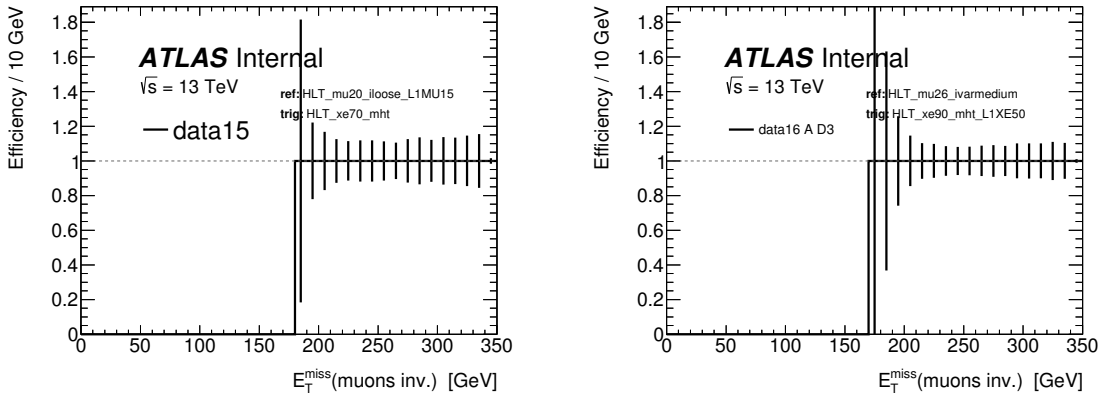


(c) 2018 B-C5



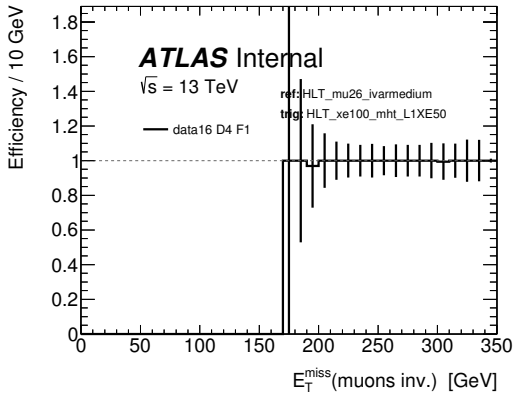
(d) 2018 C6-

Fig. 4.5 2D trigger efficiency E_T^{miss} vs $p_{\text{T}}(j_1)$ for each data taking period from 2017 to the end of 2018. Superimposed onto each plot is the proposed 2D selection that would ensure the trigger is fully efficient.

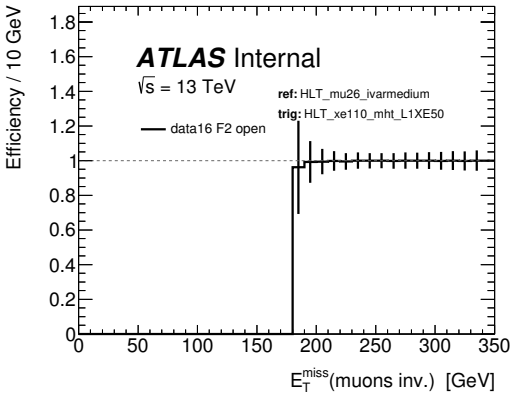


(a) 2016 A-D3

(b) 2016 A-D3

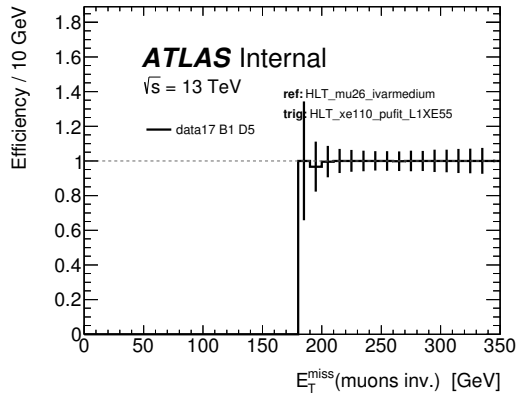


(c) 2016 D4-F1

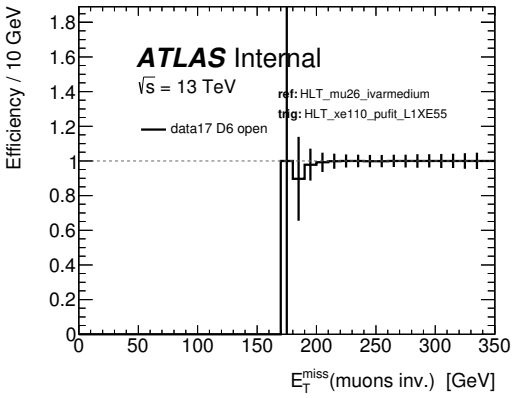


(d) 2016 F2-

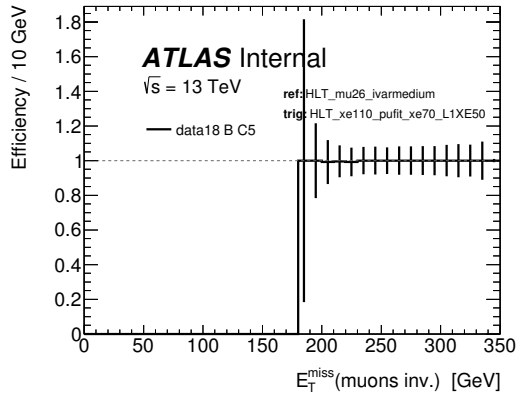
Fig. 4.6 Trigger efficiency as a function of E_T^{miss} for each data taking period from 2015 to 2016. A selection is applied here of trigPlateau > 1, $\frac{p_T(j_1)}{HT} > 0.7$ and object based E_T^{miss} significance > 7. Systematic uncertainties are not considered here.



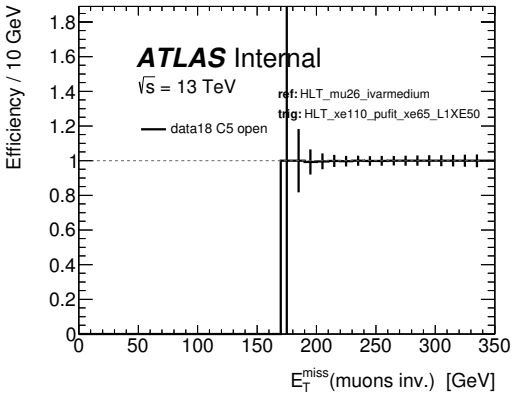
(a) 2017 B1-D5



(b) 2017 D6-



(c) 2018 B-C5



(d) 2018 C6-

Fig. 4.7 Trigger efficiency as a function of E_T^{miss} for each data taking period from 2017 to 2018. A selection is applied here of trigPlateau > 1, $\frac{p_T(j_1)}{HT} > 0.7$ and object based E_T^{miss} significance > 7. Systematic uncertainties are not considered here.

physics) and E_T^{miss} arising from mismeasurements of objects like jets and leptons. S is used to help eliminate multi-jet background events.

- Contransverse mass (m_{CT}): As described in reference [89], for two visible particles, the contransverse mass is defined as

$$\text{m}_{\text{CT}}^2(\text{v}_1,\text{v}_2) = [E_{\text{T}}(\text{v}_1) + E_{\text{T}}(\text{v}_2)]^2 - [\mathbf{p}_{\text{T}}(\text{v}_1) - \mathbf{p}_{\text{T}}(\text{v}_2)]^2, \tag{4.6}$$

where $E_{\text{T}} = \sqrt{p_{\text{T}}^2 + m^2}$ and v_1 and v_2 are the two visible particles being considered. This variable has a kinematic endpoint

$$\text{m}_{\text{CT}}^{\text{max}} = \frac{m_i^2 - m_X^2}{m_i}, \tag{4.7}$$

where i is the initial particle and X is the particle it decays into. This variable is useful when discriminating from top production backgrounds, which have an endpoint of 135 GeV.

Angular Variables

Angular variables are particularly useful when attempting to discriminate between DM signals and $Z+$jets events where the Z boson decays invisibly. This is because angular variables can take advantage of the difference in spin between the Z ($s=1$) and the ϕ/a DM mediators ($s=0$). The list of angular variables used in this analysis is as follows:

- $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$: Defined as the minimum $\Delta\phi$ between any of the leading three jets and the E_T^{miss} in the event. This variable is useful for discriminating DM signals from $t\bar{t}$ and is also used to help eliminate contamination from multi-jet background events.

- $\cos^*(b_1,b_2)$: As proposed for DM searches and described in Reference [90], this variable is defined as

$$\cos^*(b_1,b_2) = \left|\tanh\left(\frac{\Delta\eta(b_1,b_2)}{2}\right)\right|, \tag{4.8}$$

where b_1 and b_2 are the leading and sub-leading b-jets. This variables is particularly useful from discriminating $Z \to \nu\bar{\nu}$ events from DM signal events due to its sensitivity

to the spin of the mediator particle. The use of the cosine in this notation comes from the following equality used in the derivation of this variable:

$$\tanh\left(\frac{\Delta\eta(b_1,b_2)}{2}\right) = \cos\left(2\tan^{-1}\left(\exp\left(\frac{\Delta\eta(b_1,b_2)}{2}\right)\right)\right). \tag{4.9}$$

- δ^+ and δ^-: Defined as

$$\begin{aligned}\delta^- &= \Delta\phi(E_T^{\text{miss}}, j^{1-3}) - \Delta\phi(b_1,b_2)\\ \delta^+ &= \left|\Delta\phi(E_T^{\text{miss}}, j^{1-3}) + \Delta\phi(b_1,b_2) - \pi\right|.\end{aligned} \tag{4.10}$$

  Similarly to $\cos^*(b_1,b_2)$, these variables are also good at discriminating between processes mediated by particles with different spins.

Background Suppression Variables

The following is a list of other variables used to separate signal from background:

- H_{ratio}^T: The ratio between $p_T(j_1)$ and HT defined by

$$H_{\text{ratio}}^T = \frac{p_{\text{T}}(j_1)}{H_{\text{T}}}. \tag{4.11}$$

  This variable is useful particularly when discriminating against $t\bar{t}$ events and also in helping to eliminate the multi-jet background.

- m_{bb}: The invariant mass of the combined two b-jet system in the event. This variable is useful for discriminating signal models from $Z+$jets and $t\bar{t}$.

- m_{ll}: The invariant mass of the combined two lepton system. This variable is mostly used for the $Z+$jets CRs in order to select events where the leptons mostly likely originate from Z-boson decay.

- m_{T}: The transverse mass of the combined lepton-E_T^{miss} system. In regions where there is only one lepton, this variable helps to reduce the amount of fake leptons arising from the misidentification of jets which disproportionately populate low values of m_{T}.

4.4.5 Preselection

Table 4.4 details the zero and two lepton preselections used in this analysis. These preselections require that the SUSY1 filtering, as defined in Section 4.4.1, and event cleaning are in

place, the triggers are fully efficient, any lepton criteria is imposed and the QCD background has been suppressed. The zero and two lepton preselection criteria will henceforth be referred to as pre-2b-2j-0l-lowMET and pre-2b-2j-2l-lowMET respectively.

Criterion	0-lepton	2 lepton	
SUSY1 Filtering	✓	✓	
Event Cleaning	✓	✓	
Lowest unprescaled trigger	HLT_xe	HLT_mu || HLT_e	Trigger
trigPlateau		>1	Trigger
Number of signal jets		$\geq 2, \leq 3$	Signal criteria
Number of b-jets		≥ 2	Signal criteria
Number of baseline leptons	0	2	
Number of signal leptons	0	2(SF,OS)	Lepton criteria
p_{T}^{l1}	-	>27 GeV	Lepton criteria
p_{T}^{l2}	-	>20 GeV	Lepton criteria
m_{T}	-	>20 GeV	Lepton criteria
$\Delta\phi(E_T^{\mathrm{miss}}, j^{1-3})$		>0.4	QCD suppression
S		7	QCD suppression
H_{ratio}^{T}		>0.7	QCD suppression

Table 4.4 List of criteria defining the analysis zero and two lepton preselections. SFOS stands for same flavour opposite sign.

Table 4.5 shows the data and MC yields for pre-2b-2j-0l-lowMET and pre-2b-2j-2l-lowMET. The total yield from 'other' includes contributions from diboson events as well as $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$. No uncertainties are shown here as these numbers come directly from MC and data.

Modelling

The modeling of the preselections is examined for some important kinematic variables. For all plots the MC SM backgrounds have been stacked so that the total of these can be compared to data. Conversely, the contributions from two DM models have not been stacked. Due to the contributions from these signals being small, a scale factor of either 10 or 20 has been applied to them. The hatched area in the ratio plot corresponds to the total statistical error on combined MC background. For some figures auto-blinding has been implemented, blinding any data where the signal yield is over 10% of the data. This corresponds to a grey shaded area.

pre-2b-2j-0l-lowMET Figures 4.8, 4.9, 4.10, 4.11, and 4.12 show data-MC comparisons for some of the major variables in the pre-2b-2j-0l-lowMET region. Where there is a bin

Yields	pre-2b-2j-0l-lowMET	pre-2b-2j-2l-lowMET
Observed events	10686	3306
MC exp. SM events	8891.93	3276.38
MC exp. Z events	5132.18	1070.47
MC exp. W events	1441.70	0.81
MC exp. ttbar events	1760.79	1818.53
MC exp. st events	300.27	337.29
MC exp. other events	256.99	49.28
$m(\phi,\chi) = (10,1)$ GeV $, g = 1$	32.72	0.00
$m(\phi,\chi) = (100,1)$ GeV $, g = 1$	18.09	0.00

Table 4.5 Data and expected yields from MC for pre-2b-2j-0l-lowMET and pre-2b-2j-2l-lowMET.

with signal contamination is $>10\%$ of the SM background, all data in that bin and to the right has been blinded. Only statistical uncertainties are shown in these plots. There is some disagreement between the MC and the data, but generally it they are consistent to within a scale factor. This somewhat motivates the need for a more careful consideration of the major backgrounds in this analysis.

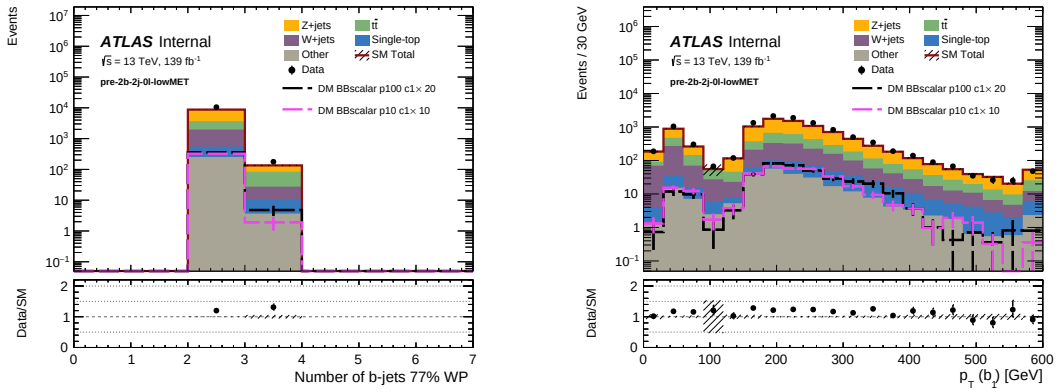


Fig. 4.8 Data MC comparisons for (a) the number of b-jets and (b) the p_T of the leading b-jet for pre-2b-2j-0l-lowMET. Uncertainties are statistical only.

pre-2b-2j-2l-lowMET Figures 4.13, 4.14, 4.15, 4.16, and 4.17 show data-MC comparisons for some of the major variables in the pre-2b-2j-2l-lowMET region. These plots are not blinded as the two lepton criteria assure orthogonality with the SRs.

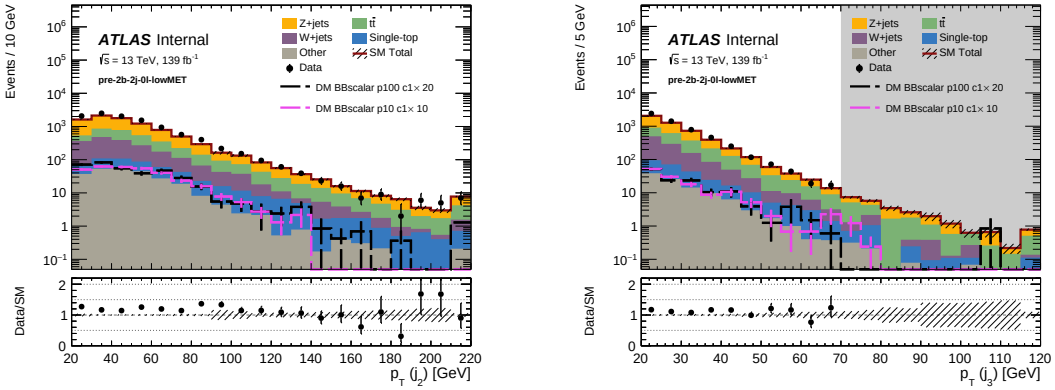


Fig. 4.9 Data MC comparisons for (a) the p_T of the second leading jet and (b) the p_T of the third leading jet for pre-2b-2j-0l-lowMET. Uncertainties are statistical only.

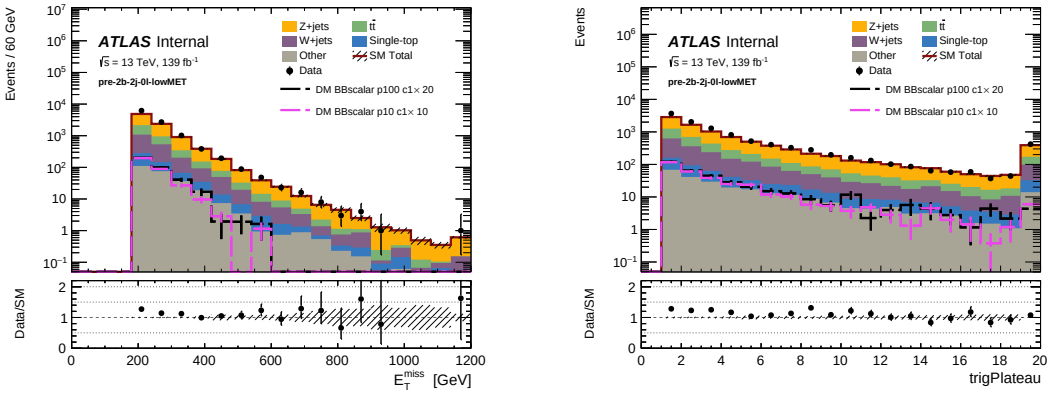


Fig. 4.10 Data MC comparisons for (a) E_T^{miss} and (b) trigPlateau for pre-2b-2j-0l-lowMET. Uncertainties are statistical only.

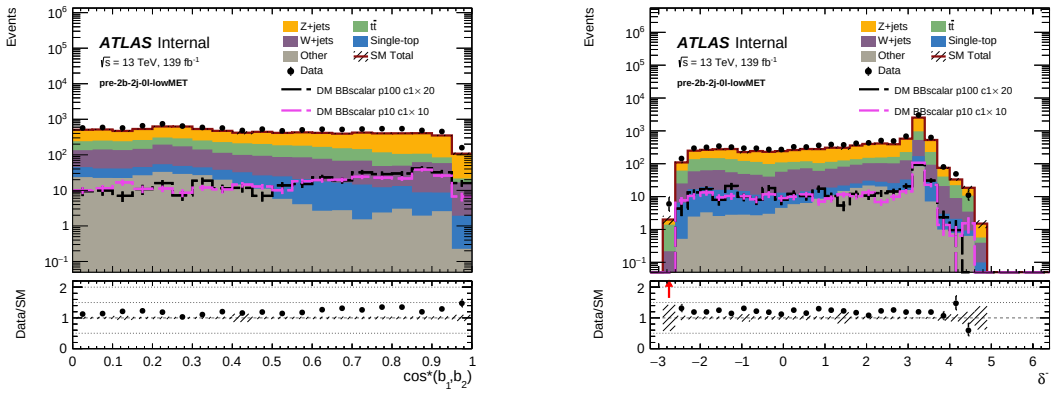


Fig. 4.11 Data MC comparisons for (a) $\cos^*(b_1, b_2)$ and (b) δ^- for pre-2b-2j-0l-lowMET. Uncertainties are statistical only.

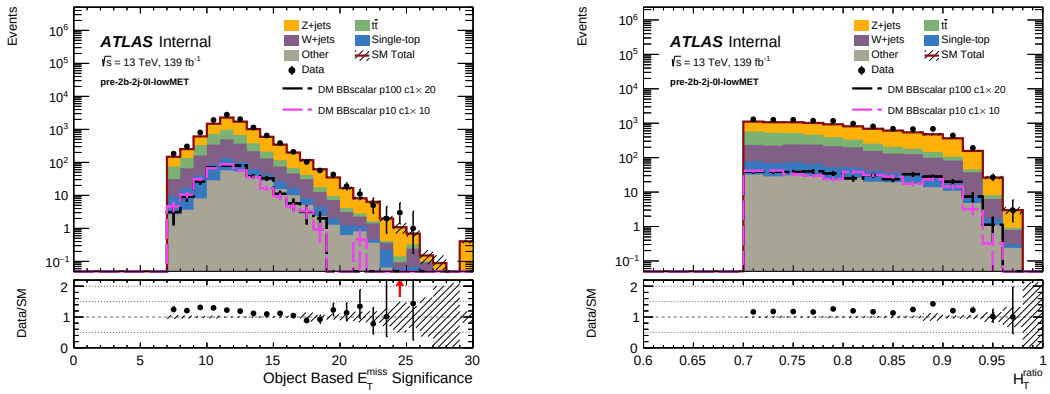


Fig. 4.12 Data MC comparisons for (a) S and (b) H^T_{ratio} for pre-2b-2j-0l-lowMET. Uncertainties are statistical only.

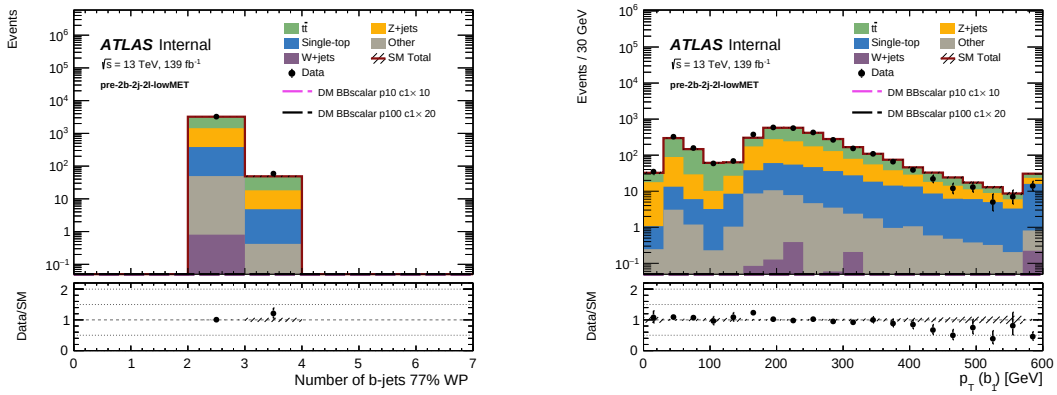


Fig. 4.13 Data MC comparisons for (a) the number of b-jets and (b) the p_T of the leading jet for pre-2b-2j-2l-lowMET. Uncertainties are statistical only.

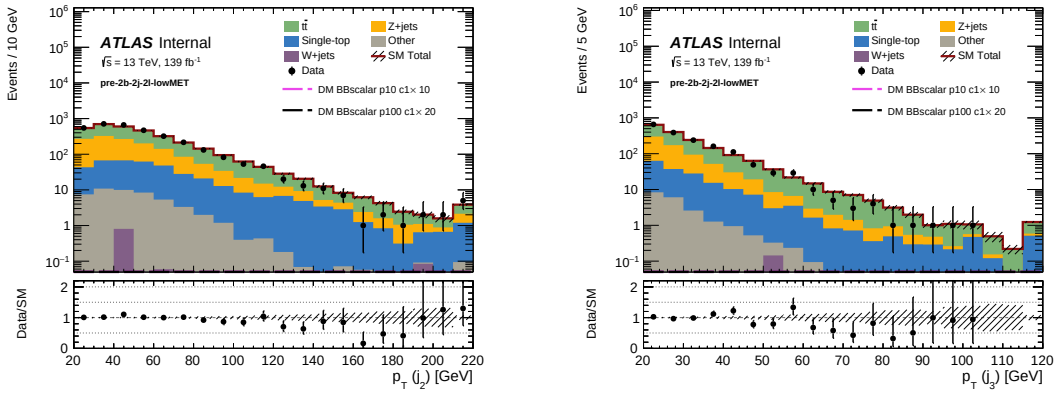


Fig. 4.14 Data MC comparisons for (a) the p_T of the second leading jet and (b) the p_T of the third leading jet for pre-2b-2j-2l-lowMET. Uncertainties are statistical only.

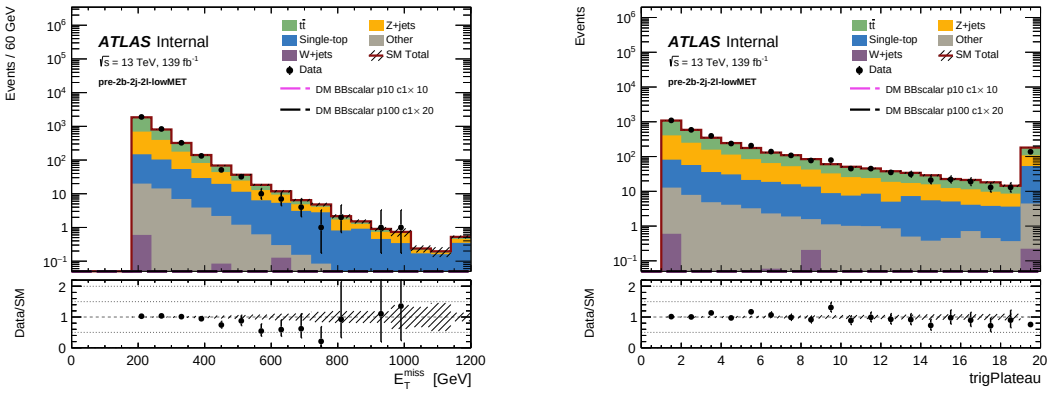


Fig. 4.15 Data MC comparisons for (a) the number of b-jets and (b) the p_T of the leading jet for pre-2b-2j-2l-lowMET. Uncertainties are statistical only.

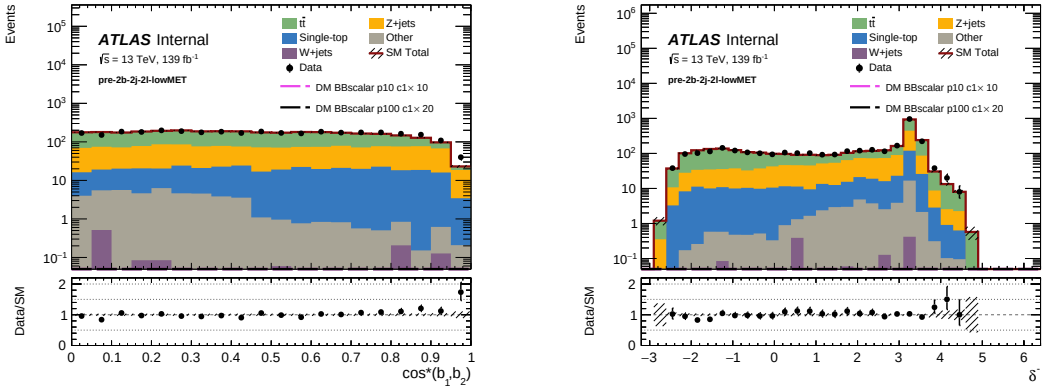


Fig. 4.16 Data MC comparisons for (a) $\cos^*(b_1, b_2)$ and (b) δ^- for pre-2b-2j-2l-lowMET. Uncertainties are statistical only.

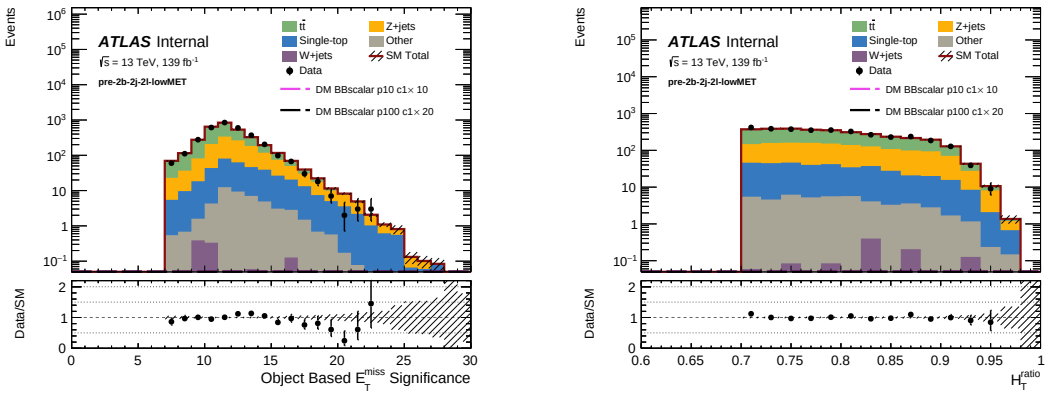


Fig. 4.17 Data MC comparisons for (a) S and (b) H^T_{ratio} for pre-2b-2j-2l-lowMET. Uncertainties are statistical only.

Chapter 5

Boosted Decision Trees and Signal Region Strategy

5.1 Boosted Decision Trees

The ATLAS Collaboration employs machine learning (ML) algorithms in many different facets of the experiment, including, but not exhausted by, flavour-tagging, the identification of most physics objects and to provide discrimination between a physics process of interest and the associated background processes, which is the focus of this chapter.

ML algorithms take input of many different variables. Each individual variable is expected to provide some degree of discrimination between the background processes and the process of interest. These variables are then combined into one more powerful discriminator via the algorithm.

The choice of ML algorithm depends on the nature of the problem. A decision was made to use Boosted Decision Trees (BDTs) as the multi-variate analysis (MVA) technique. BDTs are amongst the most common ML techniques used in high energy physics (HEP). This is partly due to the transparency of the mechanisms in place and the simplicity of the algorithms compared to other common approaches like support vector machines or neural networks. Other advantages of using BDTs are that, like other multi-variate techniques, they are good at handling multi-variate data and they can also be quick to train.

BDTs are employed in this analysis to help provide additional discrimination between DM signals and their SM backgrounds. Details of this approach are laid out in this chapter.

5.1.1 Decision Tree Learning

A Decision Tree (DT) is a series of binary criteria on variables that a data point (or an event) can either pass or fail. Each decision is a node and a 'parent' node yields two 'daughter' nodes. This is performed on the entire data set until a termination criterion is met. The choice of cut is made based on maximising the 'separation gain' between nodes. This is defined as

$$\text{separation gain} \approx \text{gain(parent node) - gain(daughter node 1)} - \text{ gain (daughter node 2)}, \tag{5.1}$$

where gain is defined as

$$\text{gain(node)} \approx p \cdot (1-p). \tag{5.2}$$

Here, p simply corresponds to the number of signal events divided by the number of background events. This definition of gain is known as the Gini Index. This is the default metric used in the Toolkit for Multi-Variate Analysis (TMVA) [91], which is a package within the ROOT software package designed specifically for ML in the context of HEP and is used for this analysis.

5.1.2 Boosting

BDTs share some of the same principles as Random Forrest techniques. A Random Forest (RF) is an ensemble technique which is used to combine the outputs from lots of different trees. One approach to generate the RF of trees is to use the Adaptive Boosting algorithm [92], which is one of the most common methods employed in HEP. This is an iterative method by which misclassified events are reweighted, or boosted, before each training. Consider a vector of variables $\vec{x}_i$ and the corresponding set of classifier outcomes, y_i, where $y = \pm 1$. The function $h_j(\mathbf{x}_i)$, where x_i is an individual data point, yields either $+1$ if the node is dominated by signal or -1 if the node is background dominated. The error on each tree, j, is given by

$$\varepsilon_j = \frac{\sum_{y_i \neq h_j(\mathbf{x}_i)} w_i^{(j)}}{\sum_{i=1}^{n} w_i^{(j)}}, \tag{5.3}$$

which is the sum of the weights, $w^(j)$, of misclassified events over the total sum of weights. ε is minimised for each tree. The corresponding score is then given by

$$\alpha_j = \beta \ln \frac{1-\varepsilon_j}{\varepsilon_j}, \tag{5.4}$$

where β is the learning rate of the BDT, which is a free parameter. The next tree then has the weights of events boosted according to the following rule

$$w_i^{(j+1)} = w_i^{(j)} e^{\alpha_j}, \tag{5.5}$$

such that the total number of events remains constant. Finally, the output of the classifier for a given event is

$$h_{fin}(\mathbf{x}_i) = \frac{1}{N_{\text{trees}}} \sum_{j=1}^{N_{\text{trees}}} \alpha_j h_j(\mathbf{x}_i). \tag{5.6}$$

There are a number of parameters which can be varied when using this type of BDT. These are:

- The total number of DTs averaged over, N_{trees}.
- The learning rate, β, as in Equation 5.4.
- The depth of the DTs (i.e. the maximum number of layers in a tree).
- The minimum number of events in a node as a percentage of the total dataset.

5.1.3 Strategy

From Section 4.4.5 there is a clear indication that the dominant background in this analysis will be Z+jets and that there will be considerable contributions from $t\bar{t}$ and W+jets also. For this reason, three separate BDTs are trained to attempt to separate these backgrounds from the DM signal models. For Z+jets, only $Z(\rightarrow \nu\bar{\nu}) +$ jets events are used in the BDT training because these represent the subset of Z+jets that is virtually indistinguishable from the signals due to the large missing energy associated with the neutrinos. Only $t\bar{t}$ events that have been generated with at least one lepton ($t\bar{t}_{\text{non-all-had}}$) events are used for training this BDT. Leptons can fake jets, meaning that there is some real missing energy coming from a $W \rightarrow l\nu$ decay, where the lepton is misidentified as a jet.

For the DM signal models that the backgrounds are trained against, a choice is also made to train on high and low mediator masses separately. Low mass mediators are defined as $m(\phi/a)[\text{GeV}] = 10, 20, 50, 100$ and high mass mediators are defined as $m(\phi/a)[\text{GeV}] = 100, 200, 300, 500$. The 100 GeV mass point is used to train both BDTs to maximise available statistics. It is expected that mass scale variables, examples of which are given in Section 4.4.4, will perform better for discriminating DM models with high mass mediators,

whilst lower mass mediators may have to rely more heavily upon angular variables to achieve any discrimination.

There are a total of 2×3 BDTs trained: one for each major background and for high and low mediator masses. Training is performed on MC only.

The selection applied for all BDT trainings is pre-2b-2j-0l-lowMET, which is defined in Section 4.4.5. This region of phase space is regarded as being narrow enough so as to only contain events where the background is difficult to discriminate from the signal, whilst still containing enough events to avoid overtraining. Overtraining is a phenomenon that occurs when there are not enough events to train on. This leads to the BDT training finding trends in the data which are really just statistical fluctuations. Although associated with a lack of statistics, overtraining can sometimes be mitigated by optimising the free parameters of the BDT. For example, increasing the minimum number of events on each node might lead to less fine-tuned, and therefore more general, results.

5.1.4 Input variables

To keep the BDT training under control, it is important to only select a subset of variables that are believed to provide good discriminating power. Variables should not be too correlated with each other for both the signal and the background samples as this means they likely provide little new information for the BDT. Figure 5.1 shows the correlation matrices for all of the candidate input variables considered for training the BDTs.

As $\cos^*(b_1, b_2)$ has no large correlations with any other variables apart from m_{bb}, these two variables are excluded from training the BDTs. This means that the discriminating power of $\cos^*(b_1, b_2)$ can be exploited in a shape fit in the final signal region (SR) and also this variable can be used to validate the agreement between data and MC of the BDT classifier responses. Variables where the only difference is whether they are calculated with two jets or two b-jets turn out to be very highly correlated with each other and so only the variables that are calculated using any jet (i.e. jets that aren't necessarily b-tagged) are considered. N_j^{good}, m_{eff} and $p_T(j_1)$ are very highly correlated to several other variables and are therefore not selected. All other variables are then used for training the BDTs. The full list of candidate input variables is given in Table 5.1 and it is stated whether or not they are subsequently used for training the BDTs.

The comparison between data and MC in pre-2b-2j-0l-lowMET is shown in Appendix A.1. The input variables are well modelled.

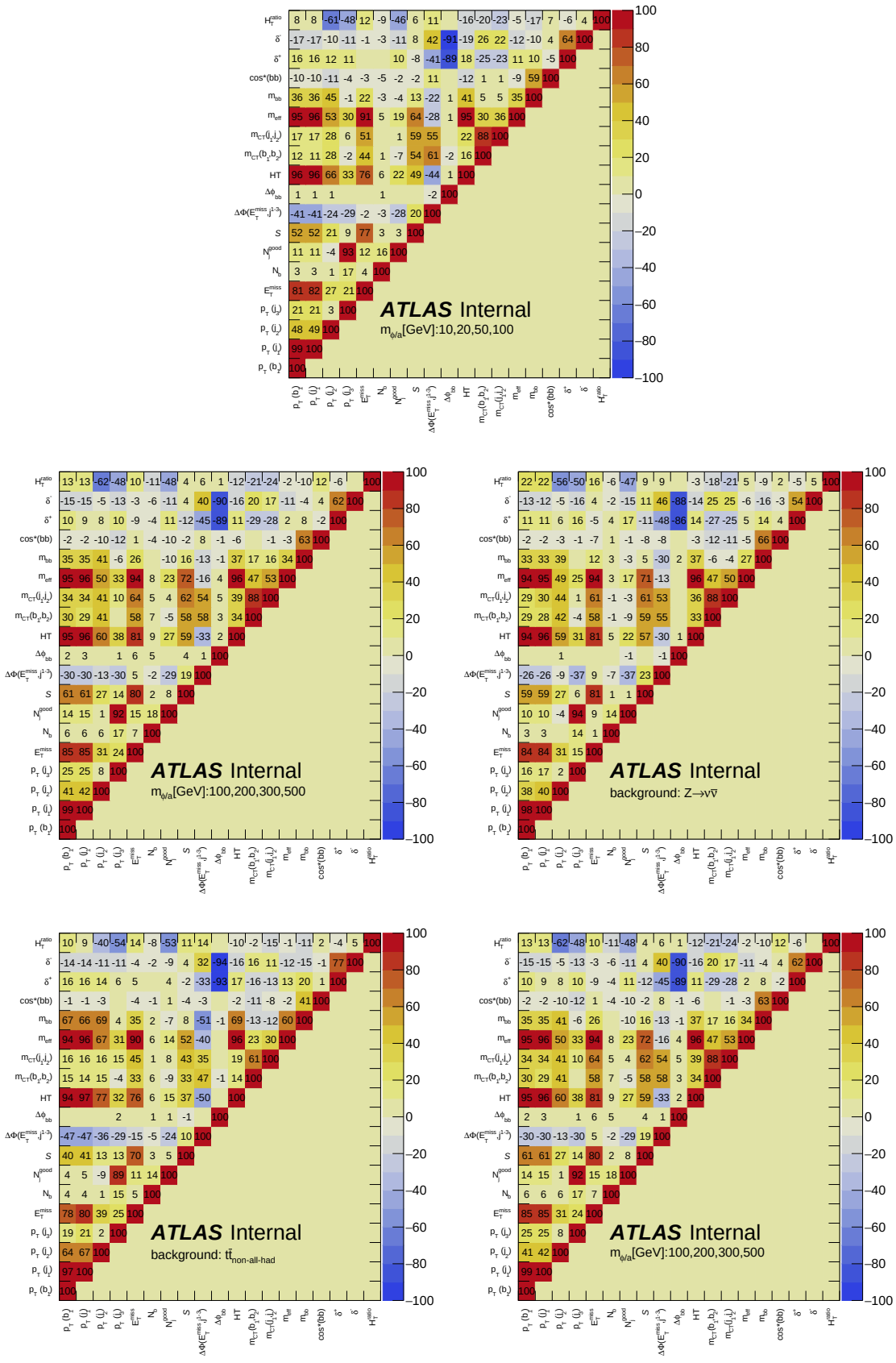


Fig. 5.1 Correlation matrices for (from top to bottom and left to right) low mass DM signals, high mass DM signals, $Z \to \nu\bar{\nu}$, $t\bar{t}_{\text{non-all-had}}$, and $W+\text{jets}$. All results are MC only.

Variable	Trained on?
$p_{\mathrm{T}}(j_1)$	
$p_{\mathrm{T}}(j_2)$	✓
$p_{\mathrm{T}}(j_3)$	✓
$p_{\mathrm{T}}(b_1)$	✓
E_T^{miss}	✓
N_j^{good}	
N_b	
object based E_T^{miss}significance (S)	✓
$\Delta\phi(E_T^{\mathrm{miss}}, j^{1-3})$	✓
$\lvert\Delta\phi_{bb}\rvert$	✓
$\mathrm{m_{CT}}(b_1, b_2)$	
$\mathrm{m_{CT}}(j_1, j_2)$	✓
HT	✓
m_{eff}	
m_{bb}	
$\cos^*(b_1, b_2)$	
δ^-	✓
δ^+	✓
H_{ratio}^T	✓

Table 5.1 Full list of candidate variables considered for BDT training. Variables that are then actually used are denoted with a tick in the right-hand column. These variables are defined in Section 4.4.4.

5.1.5 Training

Table 5.2 shows the parameters used for the BDT training. Different configurations of these parameters are chosen depending on the background as the number of events available for training each background varies. This is to minimise possible overtraining.

background	$Z \to \nu\bar{\nu}$	$t\bar{t}_{\text{non-all-had}}$	$W+\text{jets}$
N_{trees}	500	500	500
β	0.1	0.1	0.5
Tree Depth	1	1	2
Min. Node Size	10%	10%	20%

Table 5.2 BDT training configurations. High and low mass mediators use the same configuration for the same background.

Figures 5.2, 5.3 and 5.4 show the BDT classifier response for each background and for high and low mediator masses. The signal response is in blue and is generally shifted to the right, whilst the background response is in red and is generally shifted to the left. The training samples are represented by markers whilst the testing samples are represented by the solid shaded areas. The ratio plot shows the testing sample divided by the training sample for the signals whilst the bottom shows the same for the backgrounds. The metric of the level of agreement between the testing and training distributions is given by

$$\frac{\chi^2}{\text{NDOF}} = \frac{1}{\text{NDOF}} \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{5.7}$$

where O_i is the value of the observable, which is the value of the number of training sample events divided by the number of testing sample events in a given bin, i. E_i is the expectation value of the observable, which is one for all non-zero bins considered here. Finally NDOF is the number of degrees of freedom, which is the number of bins in addition to any degrees of freedom associated with the function being fitted. The function considered here is $y = 1$ and therefore there are no addition degrees of freedom associated with this.

There is reasonable agreement between the training and testing samples, with the best agreement coming from the BDTs trained on $Z \to \nu\bar{\nu}$ samples. The best discrimination comes from the BDTs trained on $t\bar{t}_{\text{non-all-had}}$ samples. For these trainings there are well defined peaks, which are clearly separated and the BDT discriminants span a larger width than the any of the others. All of the BDTs are used to define the SRs but a more in depth optimisation is performed for the $t\bar{t}_{\text{non-all-had}}$ BDTs to maximise the impact of the good separation seen in the classifier response.

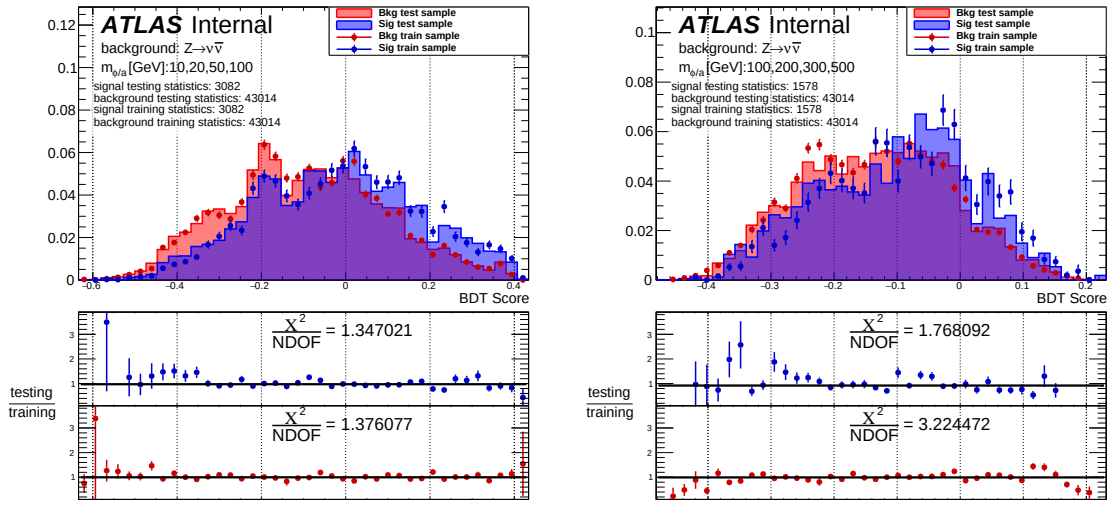


Fig. 5.2 BDT Classifier trained on $Z \to \nu\bar{\nu}$ for low mass (left) and high mass (right) signals.

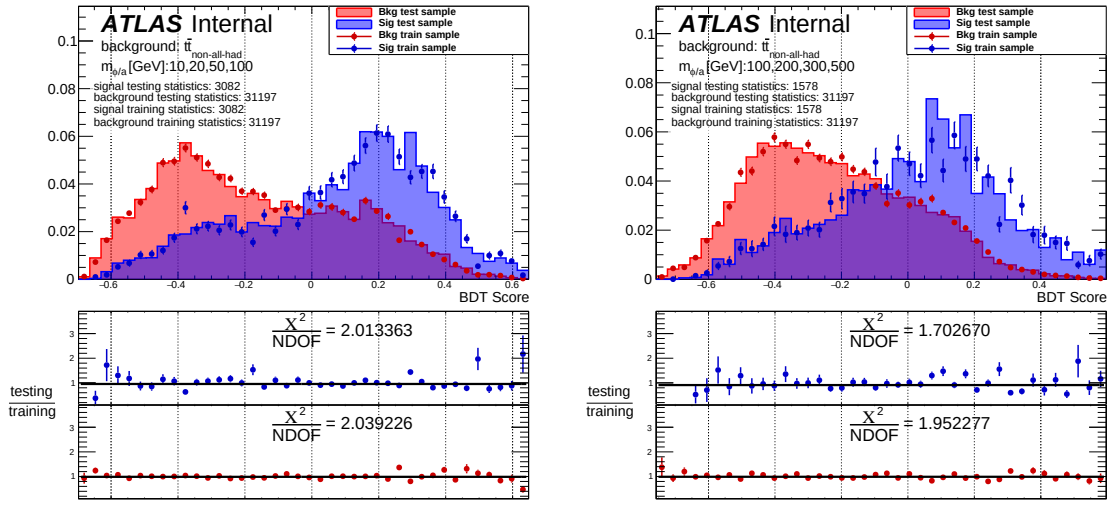


Fig. 5.3 BDT Classifier trained on $t\bar{t}_{\text{non-all-had}}$ for low mass (left) and high mass (right) signals.

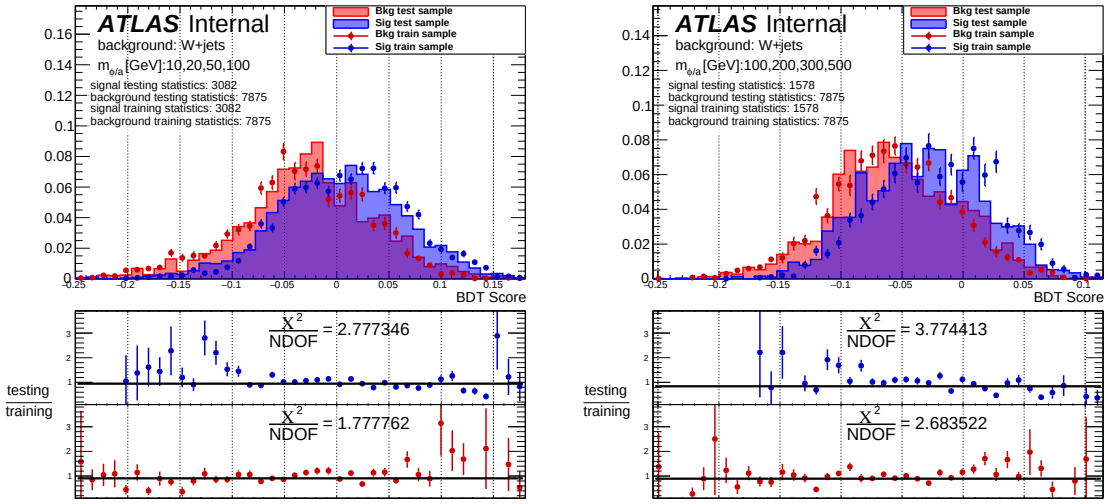


Fig. 5.4 BDT Classifier trained on $W+$jets for low mass (left) and high mass (right) signals.

The resulting BDT discriminants from the BDTs that are trained on $Z \to \nu\bar{\nu}$ against low and high mediator mass signal models are defined here as $\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{Z}_{\mathrm{highMass}}$ respectively. The resulting discriminants from the BDTs that are trained on $t\bar{t}_{\mathrm{non-all-had}}$ against low and high mass mediator signal models are defined here as $\mathrm{BDT}^{tt}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{tt}_{\mathrm{highMass}}$ respectively. Lastly, the resulting discriminant from the BDTs that are train on $W+$jets again low and high mediator mass signal models are defined here as $\mathrm{BDT}^{W}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{highMass}}$ respectively.

5.2 Signal Region Strategy

The strategy is to optimise two SRs, DM-SRB1 and DM-SRB2, which are based on cuts on BDT discriminants, the details of which are laid out in Section 5.1.3. Both SRs are designed to maximise the potential discovery and exclusion power of the DM signal models. DM-SRB1 specifically targets low mass DM mediator signal models ($m_{\phi/a}$[GeV] =10, 20, 50 and 100) and DM-SRB2 targets high mass DM mediator signal models ($m_{\phi/a}$[GeV] =200, 300 and 500).

As well as this, a cut-based SR, DM-SRA, has been optimised. The details of this are laid out in Appendix B. DM-SRA is not orthogonal to the DM-SRB1 and DM-SRB2 and so no statistical combination can be done between them. The purpose of this SR is to validate the findings of the BDT based SRs, but it is not used for the final results.

5.2.1 Optimisation

Tables 5.3 and 5.4 show limits on the production cross section of scalar DM mediators. For the sake of clarity, in this section a cross section is denoted by the symbol σ_X. This is to differentiate it from the standard symbol for a measurement of statistical significance, σ, as a relationship between these two objects is necessary for this optimisation. The figure of merit, $\frac{\sigma_X}{\sigma_X(g=1)}$, is calculated by incrementally increasing the nominal cross section (i.e. with g=1) until an expected exclusion (i.e. as calculated assuming that the data fits perfectly to the background only MC) of 2σ is achieved. This exclusion is calculated using binomial statistics. A flat systematic uncertainty of 30% is assumed in this calculation.

Given that the $t\bar{t}_{\text{non-all-had}}$ BDT trainings performed the best (i.e. distributions with the most clear separation were obtained), a decision was made to optimise with respect to these BDTs first. The cuts on these BDT distributions are in addition to the pre-2b-2j-0l-lowMET cuts. This optimisation is performed for scalar mediators only. Given that the BDT training is performed using a combination of the two models and given the shape of the signal distributions shown in Appendix A.2, an assumption is made that the result is similar for pseudo-scalar mediators.

Table 5.3 shows $\frac{\sigma_X}{\sigma_X(g=1)}$ for all scalar mediator masses with cuts on $\text{BDT}^{tt}_{\text{lowMass}}$ varying between -0.3 and 0.3 in steps of 0.1. Models with $m_\phi \leq 100$ GeV are prioritised. From the table it can be seen that a cut of either $\text{BDT}^{tt}_{\text{lowMass}} > 0$ or $\text{BDT}^{tt}_{\text{lowMass}} > 0.1$ performs optimally. In the interest of keeping as many events as possible for further optimisation and to reduce the final statistical uncertainty, a cut of $\text{BDT}^{tt}_{\text{lowMass}} > 0$ is chosen.

Table 5.4 shows $\frac{\sigma_X}{\sigma_X(g=1)}$ for all scalar mediator masses with cuts on $\text{BDT}^{tt}_{\text{highMass}}$ varying between -0.3 and 0.3 in steps of 0.1. Conversely to the previous case, models with

	$\frac{\sigma_X}{\sigma_X(g=1)}$	$m(\phi)$ [GeV]						
		10	20	50	100	200	300	500
Region	$\text{BDT}^{tt}_{\text{lowMass}} > -0.3$	113.7	126.4	148.5	206.5	477.7	957.8	5412.9
	$\text{BDT}^{tt}_{\text{lowMass}} > -0.2$	112.1	119.7	142.9	199.0	470.5	918.0	5150.2
	$\text{BDT}^{tt}_{\text{lowMass}} > -0.1$	105.5	114.2	136.6	186.6	448.5	896.2	4681.6
	$\text{BDT}^{tt}_{\text{lowMass}} > 0.0$	99.1	105.3	125.7	169.4	411.7	854.2	4398.6
	$\text{BDT}^{tt}_{\text{lowMass}} > 0.1$	96.2	110.9	129.0	165.4	406.5	795.7	3997.2
	$\text{BDT}^{tt}_{\text{lowMass}} > 0.2$	102.6	113.3	134.5	161.8	373.9	742.3	4121.2
	$\text{BDT}^{tt}_{\text{lowMass}} > 0.3$	127.2	139.6	158.5	166.2	427.9	764.5	4182.7

Table 5.3 Table of the expected exclusion for a given cut on $\text{BDT}^{tt}_{\text{lowMass}}$ between -0.3 and 0.3 in steps of 0.1. This selection is applied in addition to pre-2b-2j-0l-lowMET and the figure of merit used corresponds the scale factor applied to the nominal cross section ($\sigma_X(g=1)$) to achieve a 2σ expected significance.

$m_\phi > 100$ GeV are prioritised as this is where the final SR, DM-SRB2, is expected to outperform DM-SRB1. The signal efficiency for high mass mediator models is lower than for low mass mediator models and so, even though from the table it can be seen that a cut of $\text{BDT}^{tt}_{\text{highMass}} > 0$ does not perform quite optimally, this cut is used for the SR definition.

	$\frac{\sigma_X}{\sigma_X(g=1)}$	$m(\phi)$ [GeV]						
		10	20	50	100	200	300	500
Region	$\text{BDT}^{tt}_{\text{highMass}} > -0.3$	113.5	128.3	151.5	206.3	461.9	952.3	5434.2
	$\text{BDT}^{tt}_{\text{highMass}} > -0.2$	112.1	122.9	145.9	191.3	443.8	854.0	4926.2
	$\text{BDT}^{tt}_{\text{highMass}} > -0.1$	114.6	118.8	140.4	186.5	420.3	826.6	4402.5
	$\text{BDT}^{tt}_{\text{highMass}} > 0.0$	115.9	126.2	137.2	179.4	421.0	773.5	4038.0
	$\text{BDT}^{tt}_{\text{highMass}} > 0.1$	120.0	136.3	142.0	174.3	389.5	743.6	3696.1
	$\text{BDT}^{tt}_{\text{highMass}} > 0.2$	139.8	145.9	157.3	168.2	383.9	773.2	3548.6
	$\text{BDT}^{tt}_{\text{highMass}} > 0.3$	165.9	135.0	158.7	170.6	378.0	838.6	3274.9

Table 5.4 Table of the expected exclusion for a given cut on $\text{BDT}^{tt}_{\text{highMass}}$ between -0.3 and 0.3 in steps of 0.1. This selection is applied in addition to pre-2b-2j-0l-lowMET and the figure of merit used corresponds the scale factor applied to the nominal cross section ($\sigma_X(g=1)$) to achieve a 2σ expected significance.

The selections applied to $\text{BDT}^{Z}_{\text{lowMass}}$, $\text{BDT}^{Z}_{\text{highMass}}$, $\text{BDT}^{W}_{\text{lowMass}}$ and $\text{BDT}^{W}_{\text{highMass}}$ are chosen based on Figures 5.2 and 5.4. The selection is defined to be all events to the right of the point at which the contribution from the signal becomes larger than the contribution from the background when they are normalised to the same area. The final selections for DM-SRB1 and DM-SRB2 are given in Table 5.5.

The expected MC yields are shown in Table 5.6. Z+jets is the dominant background, as expected.

Figure 5.5 and Figure 5.6 show the n-1 plots for the BDT distributions in DM-SRB1 and DM-SRB2. An n-1 plot is a comparison plot between data and MC where the cut on the variable on the x-axis has been removed from the selection. The cut on the BDT discriminant

	DM-SRB1	DM-SRB2
preselection	pre-2b-2j-0l-lowMET	
$\mathrm{BDT}^{tt}_{\mathrm{lowMass}}$	> 0	-
$\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$	> 0	-
$\mathrm{BDT}^{W}_{\mathrm{lowMass}}$	> 0	-
$\mathrm{BDT}^{tt}_{\mathrm{highMass}}$	-	> 0
$\mathrm{BDT}^{Z}_{\mathrm{highMass}}$	-	> -0.1
$\mathrm{BDT}^{W}_{\mathrm{highMass}}$	-	> -0.05

Table 5.5 Summary of selections that define DM-SRB1 and DM-SRB2.

Yields	DM-SRB1	DM-SRB2
MC exp. SM events	950.28	1623.00
MC exp. Z events	571.70	1111.05
MC exp. ttbar events	179.45	158.90
MC exp. W events	142.41	233.46
MC exp. st events	45.50	45.05
MC exp. diboson events	9.83	71.55

Table 5.6 DM-SRB1 and DM-SRB2 MC expected yields.

being shown is removed in the SR so that the data MC agreement can be seen in a region of phase space that is very similar, but orthogonal, to the SR. The SR is blinded so as not to bias the analysis at this stage. Generally there is good agreement between data and MC.

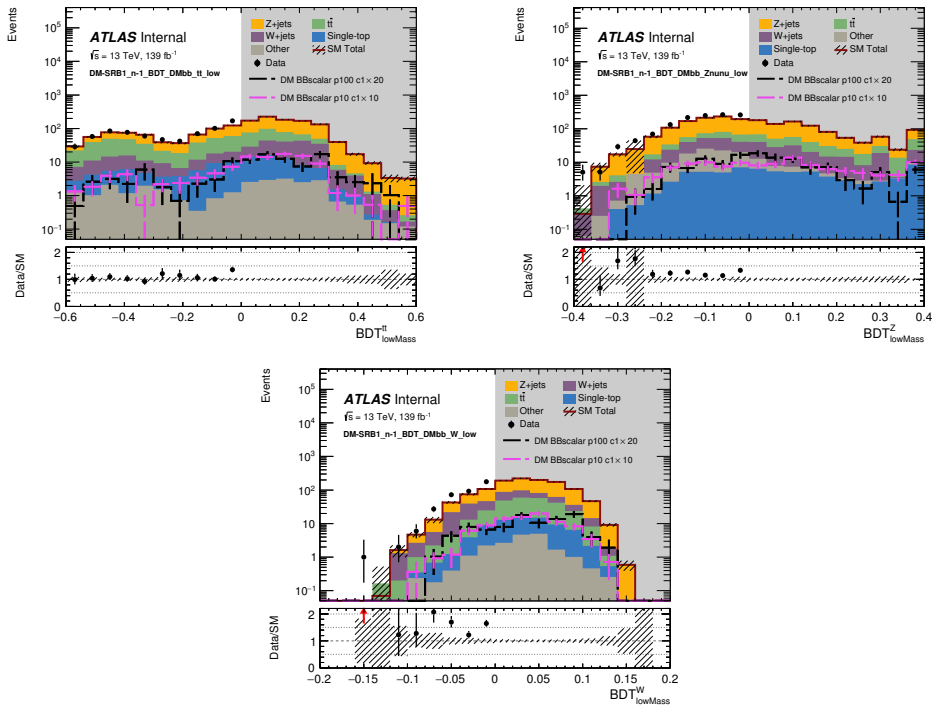


Fig. 5.5 n-1 plots for (from top to bottom and left to right) $\mathrm{BDT}^{tt}_{\mathrm{lowMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{lowMass}}$ for DM-SRB1. The grey shaded area has been blinded because this is within the SR.

Figure 5.7 and Figure 5.8 show the MC distributions of some important kinematic variables.

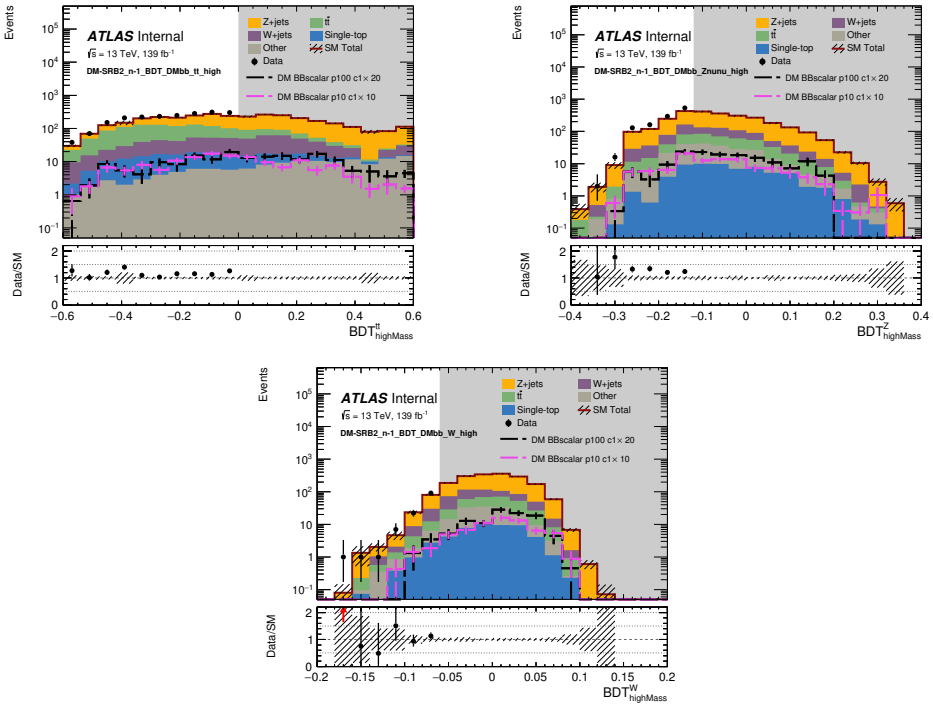


Fig. 5.6 n-1 plots for (from top to bottom and left to right) $\mathrm{BDT}^{tt}_{\mathrm{highMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{highMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{highMass}}$ for DM-SRB2. The grey shaded area has been blinded because this is within the SR.

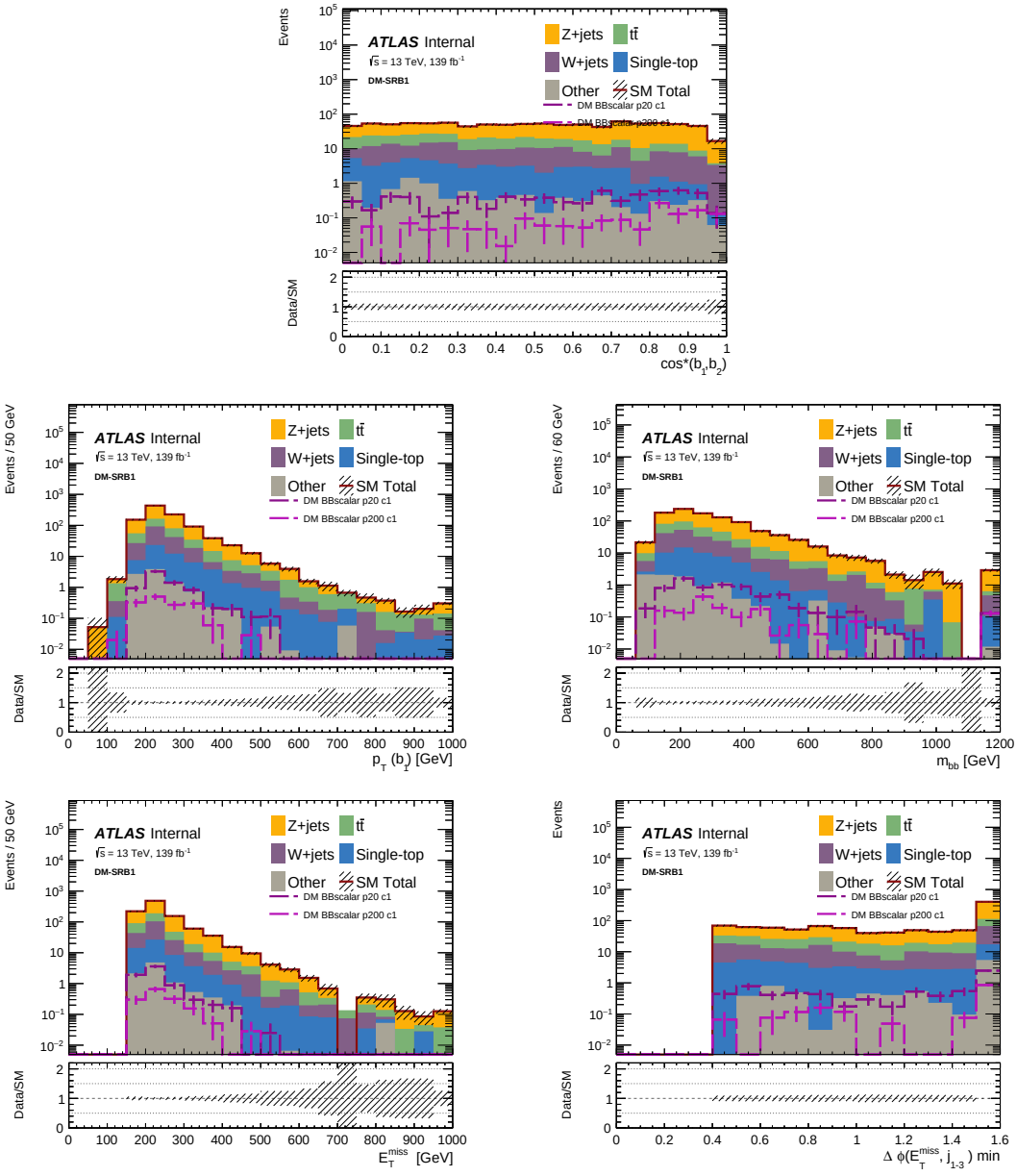


Fig. 5.7 MC only plots for (from top to bottom and left to right) $\cos^*(b_1,b_2)$, $p_T(bj_1)$, m_{bb}, E_T^{miss} and $\Delta\Phi(E_T^{\text{miss}}, j^{1-3})$ for DM-SRB1

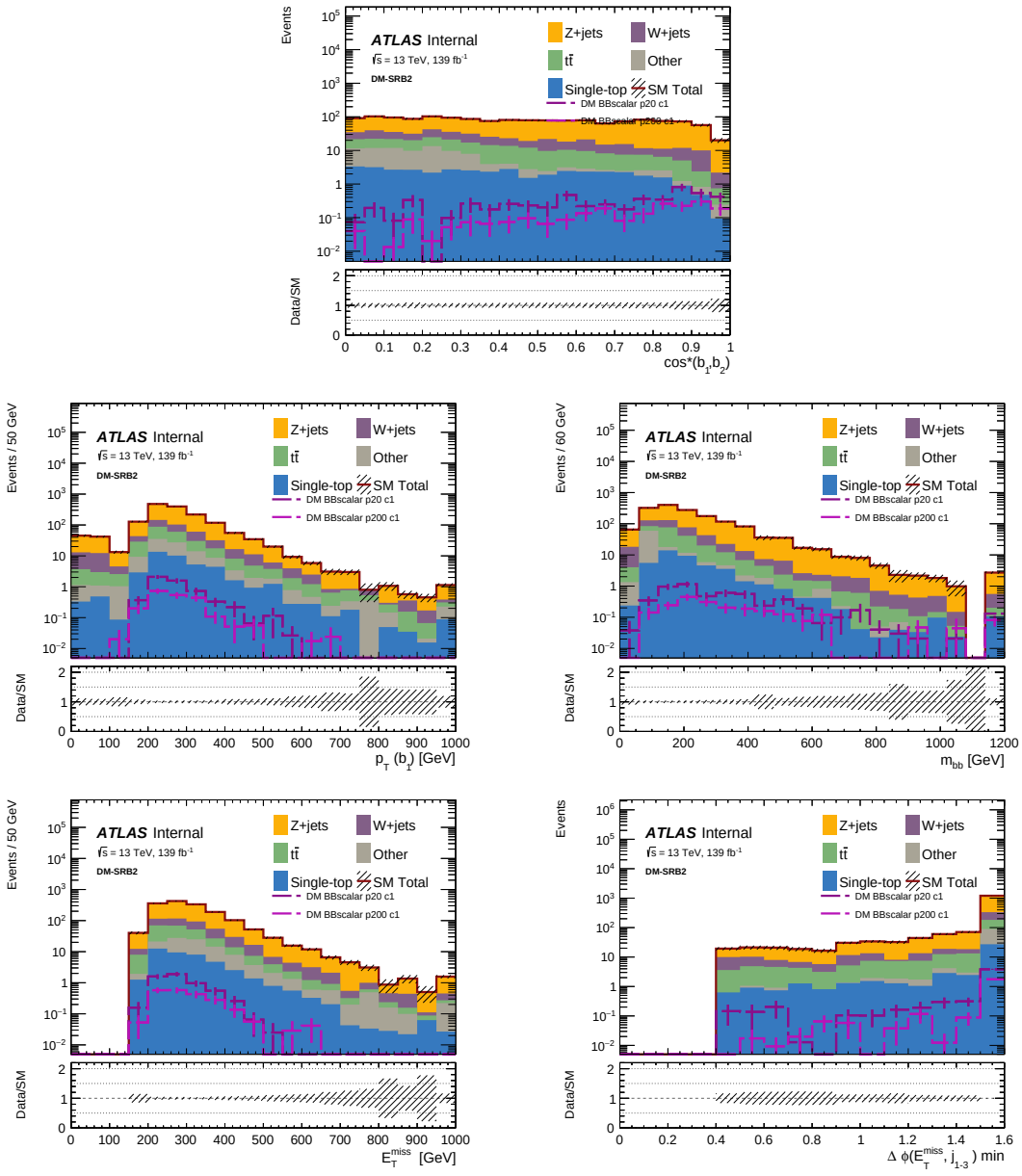


Fig. 5.8 MC only plots for (from top to bottom and left to right) $\cos^*(b_1, b_2)$, $p_T(bj_1)$, m_{bb}, E_T^{miss} and $\Delta\Phi(E_T^{\text{miss}}, j^{1-3})$ for DM-SRB2

Chapter 6

Fitting Procedure, Background Estimation and Validation

6.1 Fitting Procedure

Most backgrounds in this analysis are estimated only using MC, however the dominant background, $Z+$jets, is estimated in a semi-data-driven way by using control regions (CRs). A CR is a region of phase space which is designed to have a high purity of one type of background and should have very little contamination from any signal process. CRs should be completely orthogonal, but kinematically similar, to SRs. A CR can be used to control a background process by normalising the MC of a particular background process to data in this region. The number of events of the targeted physics process in the SR, $N_{\text{Proc}}(\text{SR},\text{est})$, is then estimated by

$$N_{\text{Proc}}(\text{SR},\text{est}) = N_{\text{Proc}}(\text{CR},\text{obs}) \cdot \frac{N_{\text{Proc}}(\text{SR},\text{MC})}{N_{\text{Proc}}(\text{CR},\text{MC})} = \mu_{\text{Proc}} \cdot N_{\text{Proc}}(\text{SR},\text{MC}), \tag{6.1}$$

where $N_{\text{Proc}}(\text{CR},\text{obs})$ is the number of events observed in the CR, $N_{\text{Proc}}(\text{CR},\text{MC})$ is the number of events in the CR as derived from MC only and $N_{\text{Proc}}(\text{SR},\text{MC})$ is number of events in the SR as derived from MC only. All variables denoted by X_{Proc} are where X only pertains to the targeted physics process. The normalisation factor (μ), which is applied to the MC of the targeted physics process in the SR, has been defined here as

$$\mu_{\text{Proc}} = \frac{N_{\text{Proc}}(\text{CR},\text{obs})}{N_{\text{Proc}}(\text{CR},\text{MC})}. \tag{6.2}$$

It is helpful to define a transfer factor (TF) for the targeted process as

$$\mathrm{TF}_{\mathrm{Proc}} = \frac{N_{\mathrm{Proc}}(\mathrm{SR},\mathrm{MC})}{N_{\mathrm{Proc}}(\mathrm{CR},\mathrm{MC})}. \tag{6.3}$$

The uncertainty on the TF is given by

$$\Delta\mathrm{TF}_{\mathrm{Proc}} = \frac{\mathrm{TF}^{\mathrm{nom}}_{\mathrm{Proc}} - \mathrm{TF}^{\mathrm{var}}_{\mathrm{Proc}}}{\mathrm{TF}^{\mathrm{nom}}_{\mathrm{Proc}}}, \tag{6.4}$$

where the $\mathrm{TF}^{\mathrm{nom}}_{\mathrm{Proc}}$ is equivalent to $\mathrm{TF}_{\mathrm{Proc}}$ as defined in Equation 6.3. $\mathrm{TF}^{\mathrm{var}}_{\mathrm{Proc}}$ derived by calculating the effect of the up and down variations on the TF due to a systematic uncertainty. $\Delta\mathrm{TF}_{\mathrm{Proc}}$ is calculated for every systematic considered in the analysis. The total uncertainty on the transfer factor is given by the summing all of the individual uncertainties in quadrature. The advantage of using a TF is that most of the effect of the systematic uncertainties is cancelled by virtue of using the ratio of MC estimates. μ values and TFs can be calculated simultaneously for multiple CRs that target different processes, but this is not necessary for this analysis where two independent fits are performed each using only one SR and one CR.

The general likelihood function, L, of an analysis that uses a CR strategy can be written as

$$\begin{aligned} L\left(n,\theta^0|\mu_{\mathrm{sig}},b,\theta\right) &= P_{\mathrm{SR}} \times P_{\mathrm{CR}} \times C_{\mathrm{syst}} \\ &= P\left(n_{\mathrm{SR}}|\lambda_{\mathrm{SR}}\left(\mu_{\mathrm{sig}},b,\theta\right)\right) \times P\left(n_{\mathrm{CR}}|\lambda_{\mathrm{CR}}\left(\mu_{\mathrm{sig}},b,\theta\right)\right) \times C_{\mathrm{syst}}\left(\theta^0,\theta\right), \end{aligned} \tag{6.5}$$

where P_{CR} and P_{SR} are the Poisson distributions of the number of events in the CR and SR respectively and C_{syst} takes into account the systematics. n_{SR} and n_{CR} are the number of observed events in the SR and CR respectively. λ_{SR} and λ_{CR} are the Poissonian prediction for the number of events in the SR and the CR respectively. They are a function of the background prediction, b, the signal strength parameter, μ_{sig}, and the nuisance parameters containing the systematic uncertainties, θ. This signal strength parameter can be set to 0 if one is performing a 'background only' fit or it is set to 1 for an 'exclusion' fit.

The final term in Equation 6.5 is the systematic term. This is a probability density function that is constructed using the product of each systematic variation modelled as a Gaussian function. The nominal value of the systematic, θ, can be varied around a central value, defined as θ^0, which can be written as

$$C_{\mathrm{syst}}\left(\theta^0,\theta\right) = \prod_{j \in Systs} G\left(\theta^0_j - \theta_j\right). \tag{6.6}$$

The final concept which is important to this analysis is that of a validation region (VR). VRs are designed to verify the background estimation as obtained from the CR. They should be kinematically very similar to the SR but orthogonal to both the CR and the SR.

This fitting procedure is implemented using `HistFitter` [93], which is a framework that is widely used within the ATLAS collaboration and has been designed specifically to perform analyses that use SRs, CRs and VRs.

6.2 Control Region Strategy

Given the background composition shown in Table 5.6, a set of two CRs has been designed for two independent fits. Both CRs target the $Z+$jets background. The first CR, DM-CRZB1, is designed to be kinematically similar to DM-SRB1 and the second, DM-CRZB2, is designed to be kinematically similar to DM-SRB2.

The CRs have two leptons, making them mutually exclusive with the SRs. They have not been designed to be orthogonal to each other (as is the case with the SRs) and so, for the final exclusion limit, no statistical combination can be performed.

These CRs are based on dilepton triggers as the main principle is to use $Z \to \ell\ell$ as a proxy for $Z \to \nu\bar{\nu}$. In the CRs, variables that are calculated using E_T^{miss} are recalculated using the p_T of the combined dilepton system as a proxy for E_T^{miss}. A mass window of 81 GeV $< m_{ll} <$ 101 GeV is applied in conjunction with $\tilde{E}_T^{\text{miss}} < 100$ GeV, where $\tilde{E}_T^{\text{miss}}$ is the missing energy calculated in the usual way without adding in the momentum of the dilepton system as defined in Section 4.3.3, to decrease contamination from $t\bar{t}$. To make the CRs more kinematically similar to the SRs, a cut of $\text{BDT}^Z_{\text{lowMass}} > 0$ is applied to define the DM-CRZB1 region and $\text{BDT}^Z_{\text{highMass}} > -0.1$ is applied which defines the DM-CRZB2 region. These cuts are identical to those used in the SRs. No additional cuts are placed on the other BDT discriminants as this will cause a considerable loss in statistics.

Table 6.1 summarises the CRs selections and Figures 6.1 and 6.3 show some comparison plots between data and MC for various kinematic variables. Generally there is good agreement between data and MC. Figures 6.2 and 6.4 shows the various BDT discriminant distributions and there is good agreement between data and MC.

Table 6.2 shows the expected yields for DM-CRZB1 and DM-CRZB2 obtained from MC. Errors have not been included on these numbers but are shown once the full fit is performed as in Table 7.1 and Table 7.5. The purity of the $Z+$jets relative to all other backgrounds is approximately 94% for both DM-CRZB1 and DM-CRZB2.

Selection	DM-CRZB1	DM-CRZB2
preselection	pre-2b-2j-2l-lowMET	
m_{ll} [GeV]	$> 81, < 101$	
$\tilde{E}_T^{\text{miss}}$[GeV]	< 100	
$\text{BDT}^{Z}_{\text{lowMass}}$	>0	-
$\text{BDT}^{Z}_{\text{highMass}}$	-	>-0.1

Table 6.1 Table of selections that define DM-CRZB1 and DM-CRZB2.

DM CR	DM-CRZB1	DM-CRZB2
Observed events	240	433
MC exp. SM events	222.60	391.53
MC exp. Z events	210.02	365.79
MC exp. W events	0.00	0.00
MC exp. ttbar events	7.99	7.99
MC exp. st events	1.90	2.88
MC exp. ttW events	0.02	0.02
MC exp. ttZ events	0.20	0.30
MC exp. ttH events	0.01	0.00
MC exp. diboson events	2.46	14.55

Table 6.2 Expected and observed yields for DM-CRZB1 and DM-CRZB2. Values shown here are pre-fit.

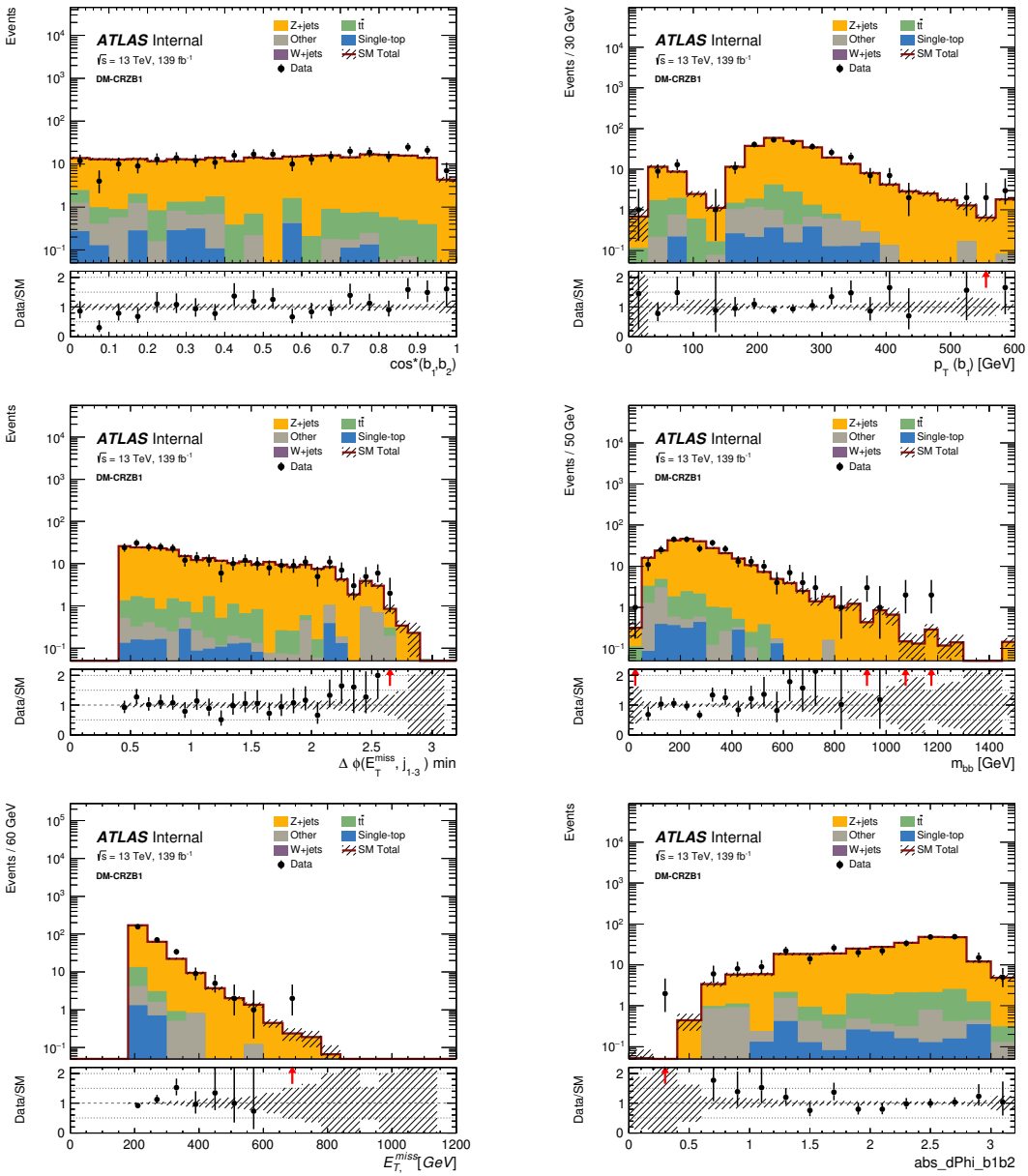


Fig. 6.1 Data/MC comparisons for (from top to bottom and left to right) $\cos^*(b_1, b_2)$, $p_T^{b_1}$, m_{bb}, $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$, E_T^{miss} and $|\Delta\phi_{bb}|$ for DM-CRZB1. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC.

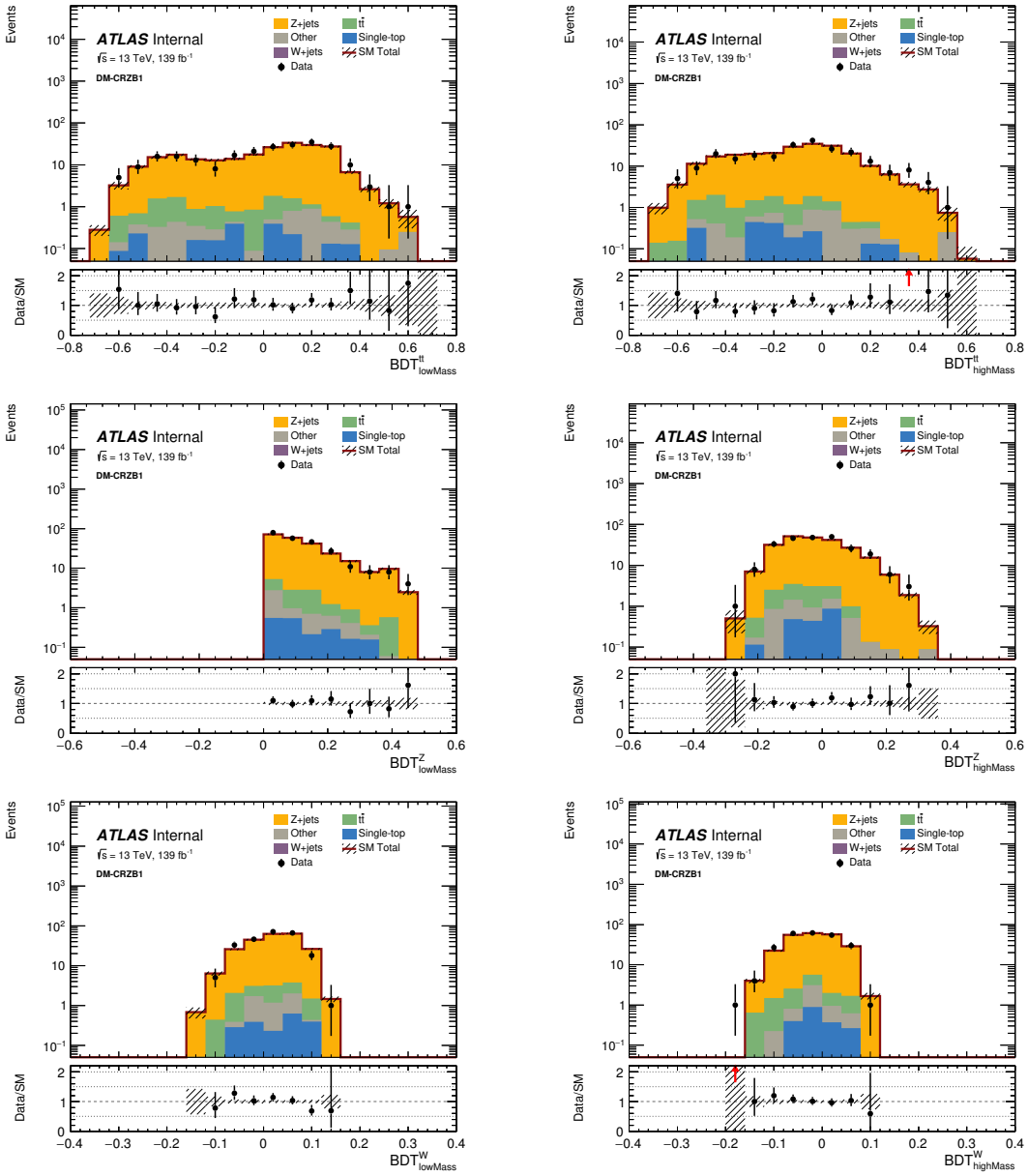


Fig. 6.2 Data/MC comparisons for (from top to bottom and left to right) $\mathrm{BDT}^{tt}_{\mathrm{lowMass}}$, $\mathrm{BDT}^{tt}_{\mathrm{highMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{highMass}}$, $\mathrm{BDT}^{W}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{highMass}}$ for DM-CRZB1. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC. A cut of $\mathrm{BDT}^{Z}_{\mathrm{lowMass}} > 0$ has been applied here.

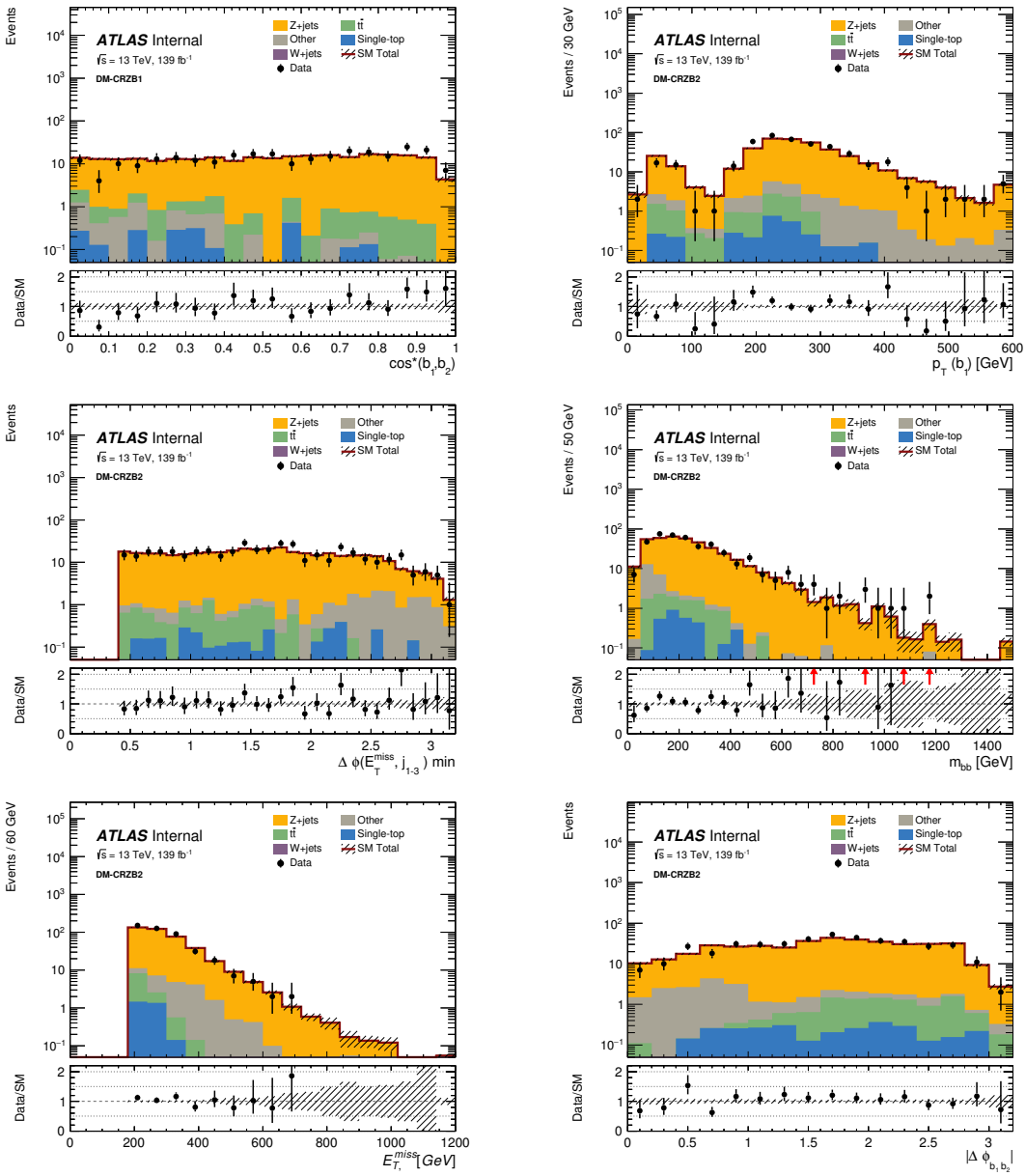


Fig. 6.3 Data/MC comparisons for (from top to bottom and left to right) $\cos^*(b_1, b_2)$, $p_T^{b_1}$, m_{bb}, $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$, $E_T^{\text{miss}}$and $|\Delta\phi_{bb}|$ for DM-CRZB2. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC.

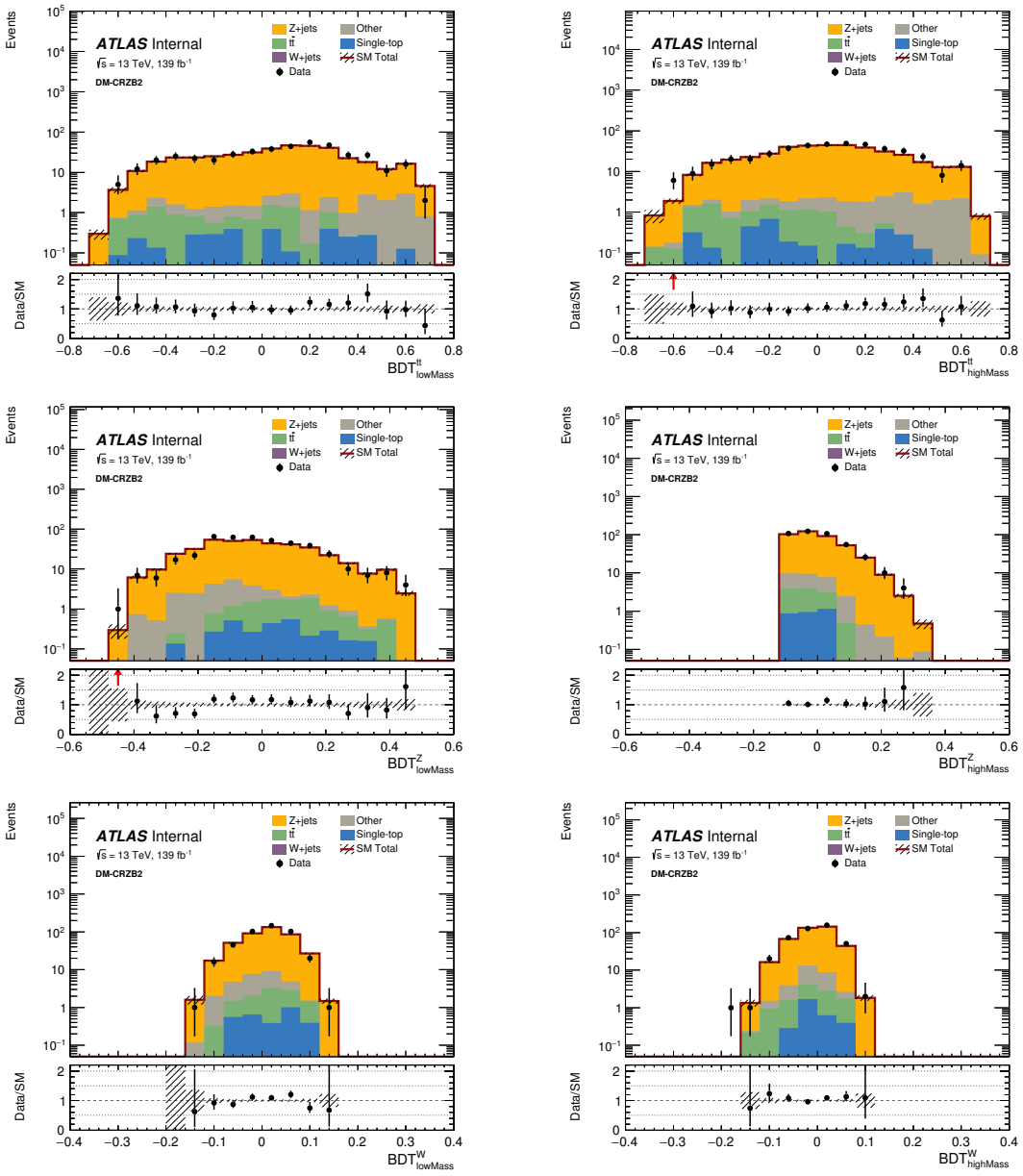


Fig. 6.4 Data/MC comparisons for (from top to bottom and left to right) $\mathrm{BDT}^{tt}_{\mathrm{lowMass}}$, $\mathrm{BDT}^{tt}_{\mathrm{highMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{highMass}}$, $\mathrm{BDT}^{W}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{highMass}}$ for DM-CRZB2. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC. A cut of $\mathrm{BDT}^{Z}_{\mathrm{highMass}} > -0.1$ has been applied here.

6.3 Background Validation

For DM-SRB1 and DM-SRB2, two VRs are defined to validate the normalisation of the $Z+$jets background from the DM CRs. DM-VRZB1 is used to validate the normalisation from DM-CRZB1 and DM-VRZB2 is designed to validate the normalisation from DM-CRZB2. Both VRs are 0-lepton regions based on the relevant SRs with an inverted selection on $\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$ or $\mathrm{BDT}^{Z}_{\mathrm{highMass}}$ for DM-SRB1 and DM-SRB2 respectively. Additional selections of $\mathrm{BDT}^{Z}_{\mathrm{lowMass}} > -0.2$ and $\mathrm{BDT}^{Z}_{\mathrm{highMass}} > -0.3$ are imposed to make the VRs as kinematically similar to the SRs as possible.

Table 6.3 shows the selections that define DM-VRZB1 and DM-VRZB2.

	DM-VRZB1	DM-VRZB2	orthogonal to SR?
preselection	pre-2b-2j-0l-lowMET		✗
BDT^{Z}_{low}	$> -0.2, < 0$	-	✓
BDT^{W}_{low}	> 0	-	✗
BDT^{tt}_{low}	> 0	-	✗
BDT^{Z}_{high}	-	$> -0.3, < -0.1$	✓
BDT^{W}_{high}	-	> -0.05	✗
BDT^{tt}_{high}	-	> -0	✗

Table 6.3 Summary of selections that define DM-VRZB1 and DM-VRZB2.

Figure 6.5 shows comparisons between data and MC for some important kinematic distributions for DM-VRZB1. The plots are pre-fit only and generally there is reasonable agreement between data and MC, however these plots are pre-fit so it is not necessarily expected that there will be perfect agreement everywhere.

Figure 6.6 shows comparisons between data and MC for some important kinematic distributions for DM-VRZB2. The plots are pre-fit only and generally there is reasonable agreement between data and MC.

Figures 6.7 and 6.8 show comparisons between data and MC of the BDT discriminants for DM-VRZB1 and DM-VRZB2 respectively. There is reasonable agreement between data and MC.

Table 6.2 shows the expected yields for DM-VRZB1 and DM-VRZB2 obtained from MC. Errors have not been included on these numbers but are shown once the full fit is performed as in Chapter 7.

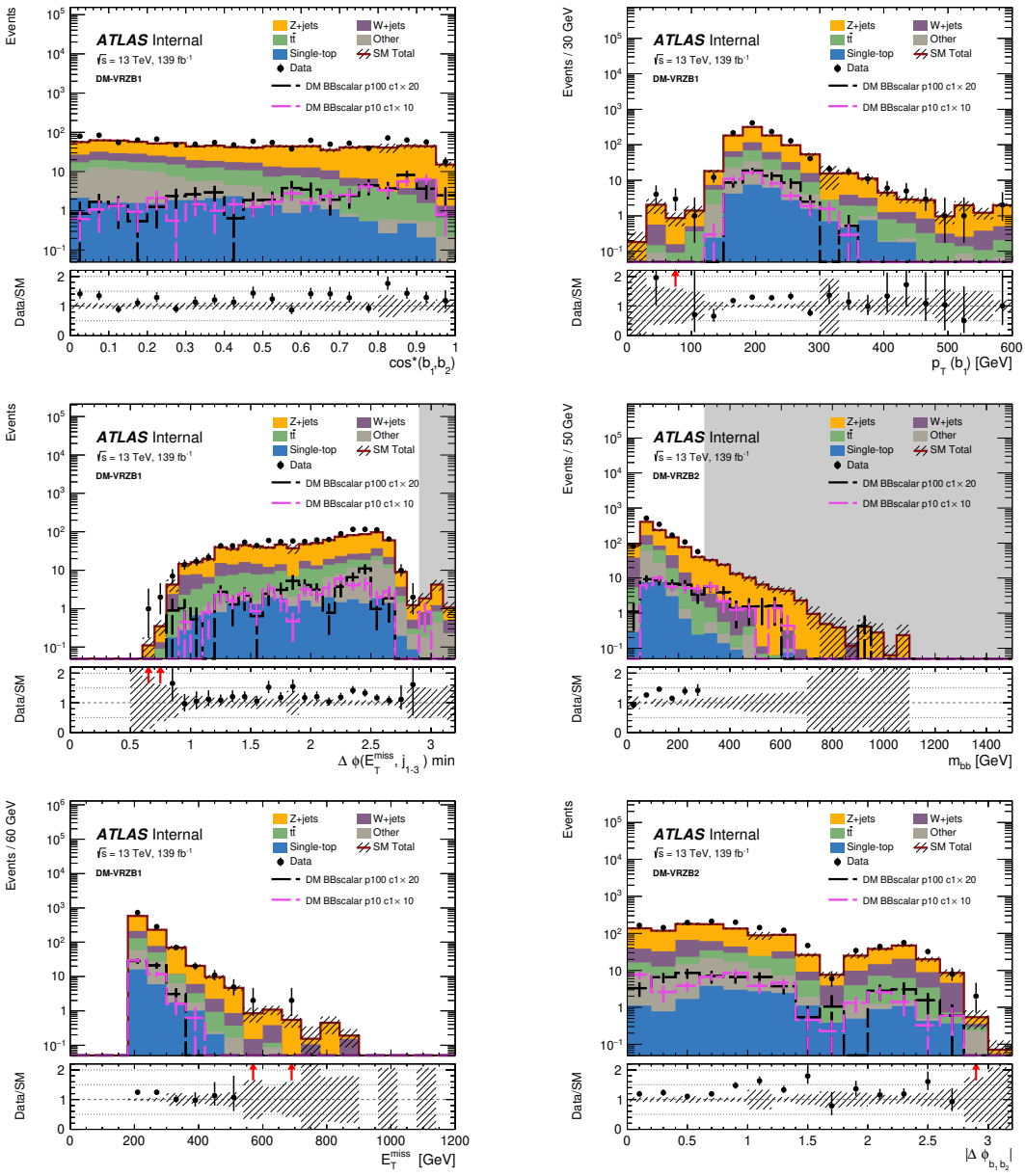


Fig. 6.5 Data/MC comparisons for (from top to bottom and left to right) $\cos^*(b_1,b_2)$, $p_T^{b_1}$, m_{bb}, $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$, E_T^{miss} and $|\Delta\phi_{bb}|$ for DM-VRZB1. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC.

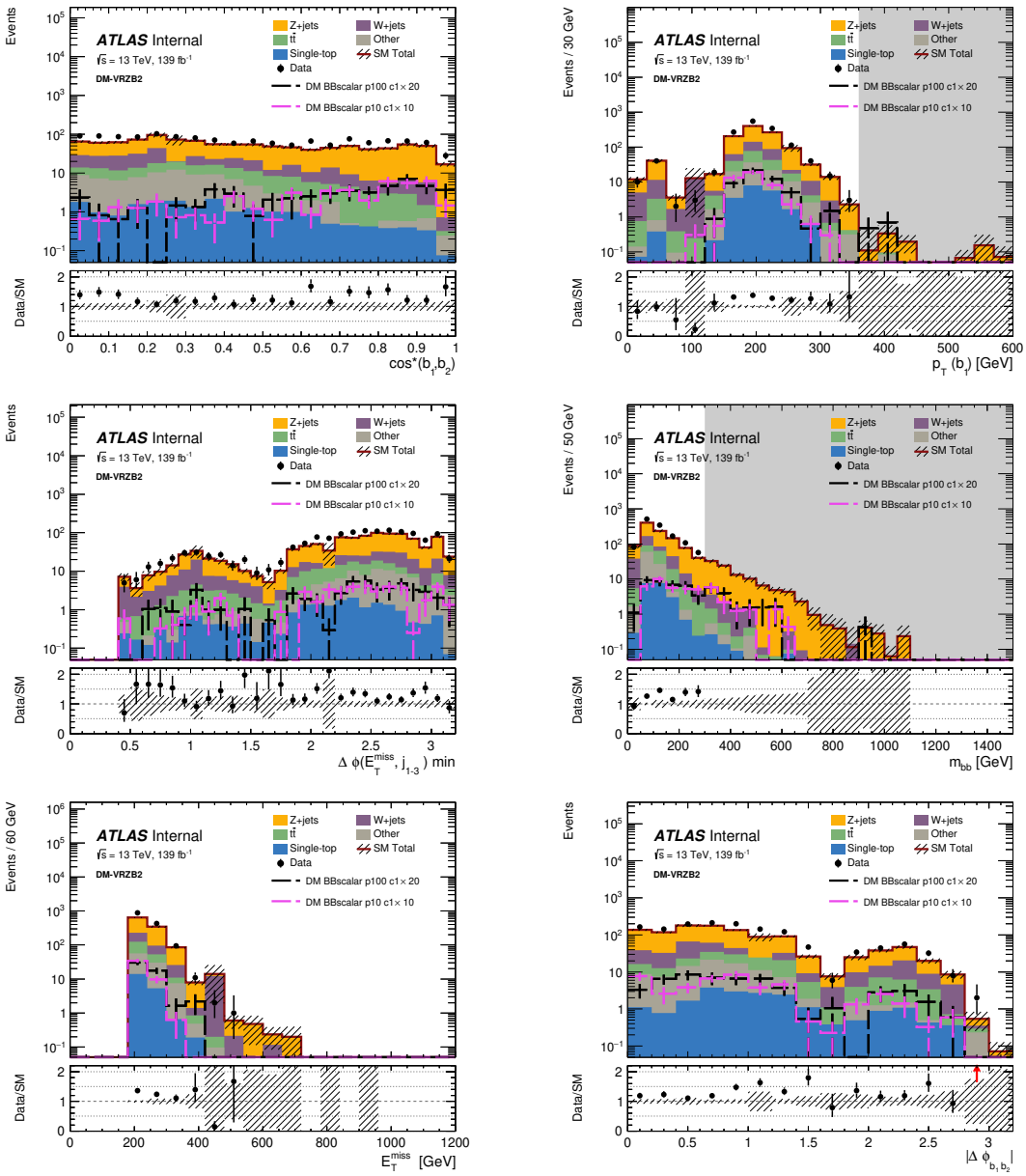


Fig. 6.6 Data/MC comparisons for (from top to bottom and left to right) $\cos^*(b_1, b_2)$, $p_T^{b_1}$, m_{bb}, $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$, E_T^{miss} and $|\Delta\phi_{bb}|$ for DM-VRZB2. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC.

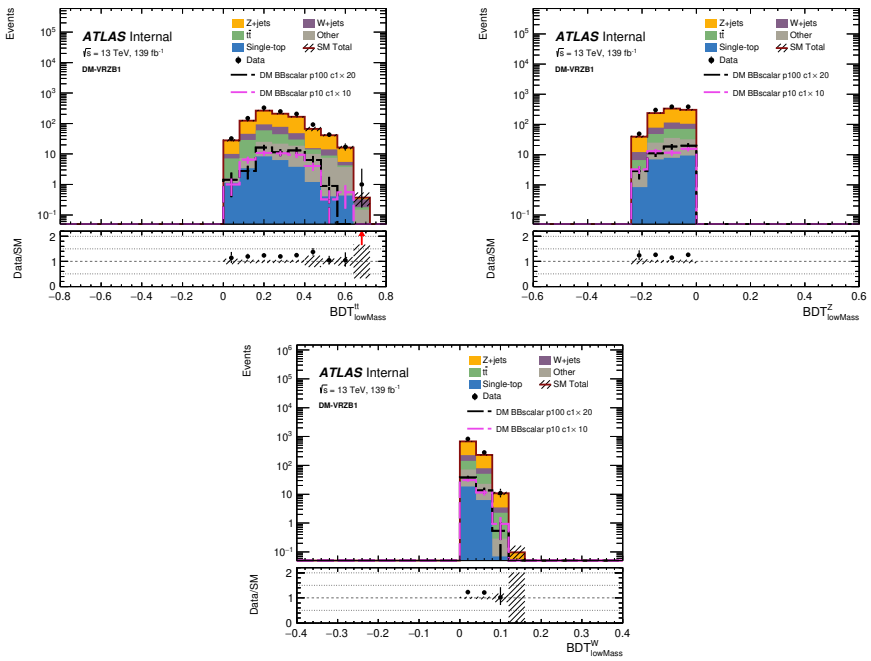


Fig. 6.7 Data/MC comparisons for (from top to bottom and left to right) $\mathrm{BDT}^{tt}_{\mathrm{lowMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{lowMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{lowMass}}$ for DM-VRZB1. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC. The selections applied are as shown in Table 6.3.

DM VR	DM-VRZB1	DM-VRZB2
Observed events	1120	1408
MC exp. SM events	898.84	1083.95
MC exp. Z events	605.54	731.88
MC exp. ttbar events	101.12	95.49
MC exp. W events	114.39	174.17
MC exp. st events	24.98	20.26
MC exp. diboson events	51.91	61.55
$m(\phi,\chi) = (10,1)$	4.46	4.33
$m(\phi,\chi) = (100,1)$	2.60	2.62

Table 6.4 Expected and observed yields for DM-VRZB1 and DM-VRZB2. Values shown here are pre-fit.

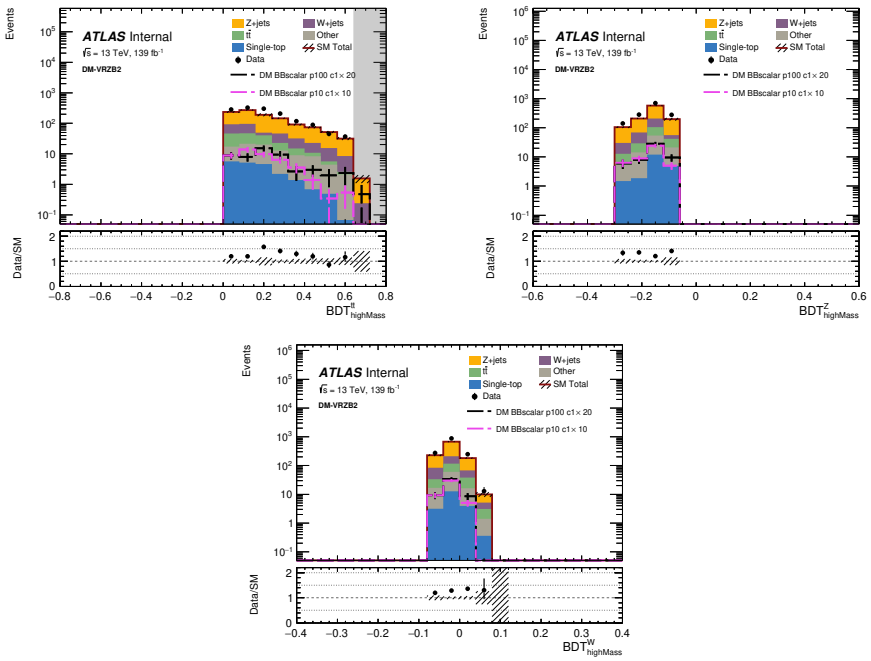


Fig. 6.8 Data/MC comparisons for (from top to bottom and left to right) $\mathrm{BDT}^{tt}_{\mathrm{highMass}}$, $\mathrm{BDT}^{Z}_{\mathrm{highMass}}$ and $\mathrm{BDT}^{W}_{\mathrm{highMass}}$ for DM-VRZB2. The error bars represent statistical uncertainties only and the hatched area in the ratio plot is the statistical uncertainty on the MC. The selections applied are as shown in Table 6.3.

Chapter 7

Results and Statistical Interpretation

The fitting procedure is as described in Section 6.1. Two independent likelihood fits are performed. The first is attempting to estimate the contribution of the $Z+$jets background in DM-SRB1, which is targeting low mass DM mediator models, using the CR, DM-CRZB1. The second fit is also targeting the $Z+$jets background, but in DM-SRB2 and using DM-CRZB2. DM-SRB2 targets high mass DM mediator models.

Separate background-only and exclusion fits are performed. The CRs are used to obtain normalisation factors ($\mu^{\text{bkg}}_{\text{region}}$). The VRs are used to show how well the data and MC agree in regions that are kinematically similar to the signals regions (SRs). The SRs do not contribute to the background only fit and are shown only as blinded VRs. In the exclusion fit, the SRs additionally constrain the fit and an additional normalisation factor for the signal strength is added (μ_{sig}). The exclusion fits are performed on five SR bins of the variable $\cos^*(b_1, b_2)$ (the bin limits are as follows $[0, 0.2, 0.4, 0.6, 0.8, 1]$).

Where the data is blinded, a place-holder for the data in the SRs is used. This is just the rounded value of the total pre-fit SM MC.

The full details of the systematic uncertainties considered in this analysis can be found in Appendix C. The systematic uncertainties can be split into two main categories; modelling uncertainties and detector uncertainties. Detector related systematics arise from the mismodelling of objects such as electrons, jets and E_T^{miss}. Modelling uncertainties arise from the choices of scales (e.g. factorisation and normalisation scales) when performing MC calculations. These uncertainties are generally symmetrised and assumed to be a Gaussian function.

7.1 DM-SRB1 Fit Results

7.1.1 Background Only Fit Results

For DM-SRB1, the region targeting low DM mediator masses ($\leq$ 100 GeV), a single CR for the $Z+$jets background is used. A value for μ is obtained for the $Z+$jets background, μ^Z_{B1}, from DM-CRZB1, where DM-CRZB1 is as defined in Section 6.2. DM-VRZB1, defined in Section 6.3, is used to test the validity of this background estimation in a region that is kinematically similar to the SR. The normalisation factor obtained in the DM-CRZB1 is

$$\mu^Z_{\text{B1}} = 1.08 \pm 0.08. \tag{7.1}$$

Table 7.1 shows the pre-fit and post-fit yields for DM-CRZB1, DM-VRZB1 and DM-SRB1. The systematic breakdown is given in Table 7.2.

Figure 7.1 shows the nuisance parameter (NP) pull plot. NP pulls are obtained by varying the systematic around its central value and are defined as $(\theta - \theta^0)/\Delta\theta$. $\Delta\theta$ is the uncertainty on the NP after the fit. These are the α terms and they represent the change in the $C_{\text{syst}}\left(\theta^0, \theta\right)$ term of Equation 6.5 and are shown on the bottom x-axis in black. The blue data points correspond to the γ terms which are Poissonian uncertainties corresponding to the μ values and the luminosity (Lumi).

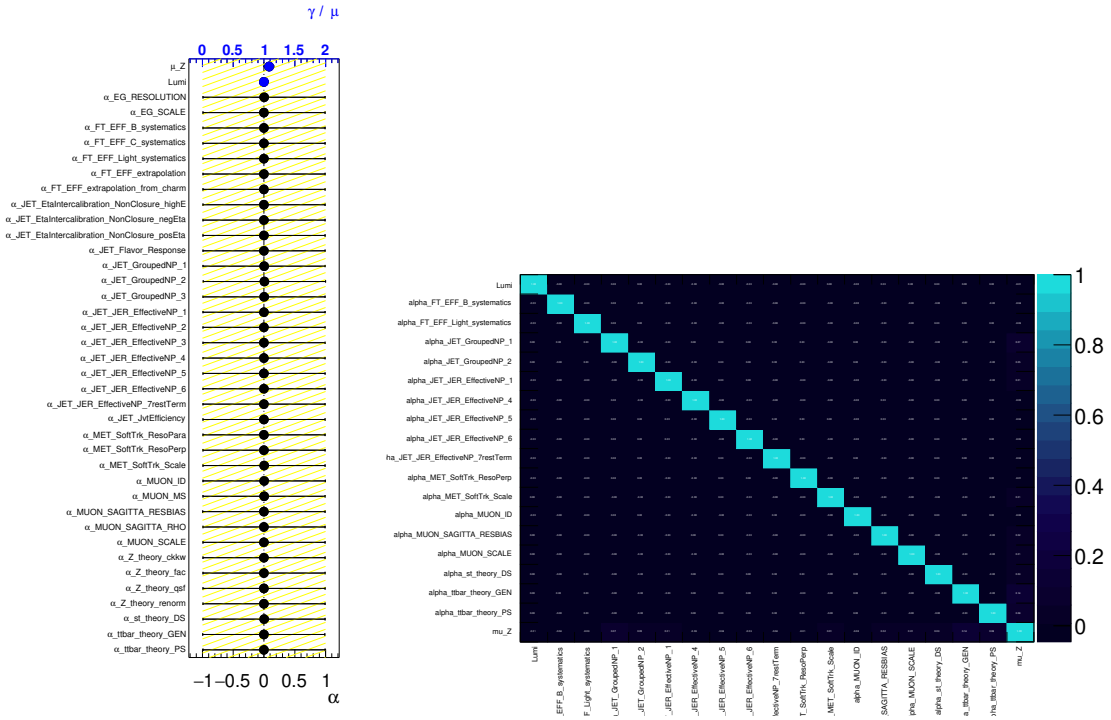
A series of pre-fit and post-fit distributions in the CR, VR and SR can be found in Figure 7.2.

	DM-CRZB1	DM-VRZB1	DM-SRB1
Observed events	240	1120	950
Fitted bkg events	240.05 ± 15.49	949.10 ± 141.76	997.79 ± 143.86
Fitted Z events	227.47 ± 15.83	655.81 ± 119.91	619.15 ± 101.63
Fitted W events	0.00 ± 0.00	114.38 ± 36.87	142.38 ± 40.29
Fitted ttbar events	7.99 ± 3.04	101.14 ± 38.18	179.57 ± 88.12
Fitted st events	1.90 ± 0.74	24.97 ± 17.96	45.47 ± 36.41
Fitted ttW events	0.02 ± 0.02	0.06 ± 0.02	0.07 ± 0.02
Fitted ttZ events	0.20 ± 0.03	0.27 ± 0.18	0.40 ± 0.13
Fitted ttH events	0.01 ± 0.00	0.56 ± 0.18	0.92 ± 0.24
Fitted diboson events	2.46 ± 0.35	51.91 ± 4.99	9.83 ± 2.19
MC exp. SM events	222.60	898.84	950.28
MC exp. Z events	210.02	605.54	571.70
MC exp. W events	0.00	114.39	142.41
MC exp. ttbar events	7.99	101.12	179.45
MC exp. st events	1.90	24.98	45.50
MC exp. ttW events	0.02	0.06	0.07
MC exp. ttZ events	0.20	0.27	0.40
MC exp. ttH events	0.01	0.56	0.92
MC exp. diboson events	2.46	51.91	9.83

Table 7.1 Fit results for one bin DM-CRZB1, DM-VRZB1 and DM-SRB1. The results are obtained from the CRs using the background-only fit. Nominal MC expectations (i.e. pre-fit) are given for comparison. The errors shown are the statistical plus systematic uncertainties.

Uncertainty of channel	DM-CRZB1	DM-VRZB1	DM-SRB1
Total background expectation	240.05	949.10	997.79
Total statistical ($\sqrt{N_{\text{exp}}}$)	±15.49	±30.81	±31.59
Total background systematic	±15.49 [6.45%]	±141.76 [14.94%]	±143.86 [14.42%]
mu_Z	±15.83 [6.6%]	±45.64 [4.8%]	±43.09 [4.3%]
alpha_ttbar_theory_GEN	±2.22 [0.93%]	±22.95 [2.4%]	±59.41 [6.0%]
alpha_ttbar_theory_PS	±1.20 [0.50%]	±26.85 [2.8%]	±60.60 [6.1%]
alpha_JET_GroupedNP_1	±1.09 [0.45%]	±21.49 [2.3%]	±36.61 [3.7%]
alpha_JET_GroupedNP_2	±0.76 [0.32%]	±21.44 [2.3%]	±29.57 [3.0%]
alpha_JET_JER_EffectiveNP_5	±0.71 [0.29%]	±11.44 [1.2%]	±7.70 [0.77%]
alpha_FT_EFF_B_systematics	±0.64 [0.27%]	±19.36 [2.0%]	±11.38 [1.1%]
alpha_JET_JER_EffectiveNP_7restTerm	±0.59 [0.25%]	±23.89 [2.5%]	±12.29 [1.2%]
alpha_JET_JER_EffectiveNP_6	±0.47 [0.19%]	±18.20 [1.9%]	±20.71 [2.1%]
alpha_st_theory_DS	±0.46 [0.19%]	±17.46 [1.8%]	±36.14 [3.6%]
alpha_JET_JER_EffectiveNP_4	±0.43 [0.18%]	±31.83 [3.4%]	±6.42 [0.64%]
alpha_FT_EFF_Light_systematics	±0.40 [0.16%]	±5.23 [0.55%]	±7.64 [0.77%]
alpha_MUON_ID	±0.37 [0.15%]	±0.07 [0.01%]	±1.12 [0.11%]
alpha_MUON_SAGITTA_RESBIAS	±0.29 [0.12%]	±0.34 [0.04%]	±0.18 [0.02%]
alpha_MUON_SCALE	±0.23 [0.09%]	±0.79 [0.08%]	±1.16 [0.12%]
alpha_MET_SoftTrk_ResoPerp	±0.21 [0.09%]	±2.92 [0.31%]	±0.35 [0.04%]
Lumi	±0.21 [0.09%]	±4.96 [0.52%]	±6.41 [0.64%]
alpha_MET_SoftTrk_Scale	±0.20 [0.08%]	±10.58 [1.1%]	±2.73 [0.27%]
alpha_JET_JER_EffectiveNP_1	±0.17 [0.07%]	±43.13 [4.5%]	±7.42 [0.74%]
alpha_JET_GroupedNP_3	±0.11 [0.05%]	±1.24 [0.13%]	±1.00 [0.10%]
alpha_JET_JER_EffectiveNP_3	±0.10 [0.04%]	±17.82 [1.9%]	±5.30 [0.53%]
alpha_JET_Flavor_Response	±0.10 [0.04%]	±1.29 [0.14%]	±0.83 [0.08%]
alpha_MET_SoftTrk_ResoPara	±0.09 [0.04%]	±2.71 [0.29%]	±2.96 [0.30%]
alpha_FT_EFF_C_systematics	±0.09 [0.04%]	±10.84 [1.1%]	±6.51 [0.65%]
alpha_MUON_MS	±0.08 [0.04%]	±1.61 [0.17%]	±1.48 [0.15%]
alpha_EG_RESOLUTION	±0.06 [0.03%]	±0.96 [0.10%]	±0.40 [0.04%]
alpha_JET_JER_EffectiveNP_2	±0.03 [0.01%]	±37.11 [3.9%]	±16.87 [1.7%]
alpha_JET_JvtEfficiency	±0.02 [0.01%]	±2.43 [0.26%]	±5.19 [0.52%]
alpha_EG_SCALE	±0.01 [0.00%]	±1.30 [0.14%]	±1.01 [0.10%]
alpha_JET_EtaIntercalibration_NonClosure_negEta	±0.00 [0.00%]	±0.27 [0.03%]	±0.55 [0.06%]
alpha_MUON_SAGITTA_RHO	±0.00 [0.00%]	±0.09 [0.01%]	±0.00 [0.00%]
gamma_stat_DMVRZB1_cosbb_bin_3	±0.00 [0.00%]	±8.80 [0.93%]	±0.00 [0.00%]
alpha_W_theory_ckkw	±0.00 [0.00%]	±8.30 [0.87%]	±10.68 [1.1%]
alpha_FT_EFF_extrapolation	±0.00 [0.00%]	±0.00 [0.00%]	±0.00 [0.00%]
alpha_W_theory_qsf	±0.00 [0.00%]	±13.66 [1.4%]	±11.84 [1.2%]
alpha_Z_theory_renorm	±0.00 [0.00%]	±89.21 [9.4%]	±80.85 [8.1%]
alpha_FT_EFF_extrapolation_from_charm	±0.00 [0.00%]	±0.00 [0.00%]	±0.00 [0.00%]
gamma_stat_DMVRZB1_cosbb_bin_4	±0.00 [0.00%]	±14.76 [1.6%]	±0.00 [0.00%]
alpha_Z_theory_ckkw	±0.00 [0.00%]	±14.08 [1.5%]	±19.67 [2.0%]
alpha_Z_theory_qsf	±0.00 [0.00%]	±21.84 [2.3%]	±15.14 [1.5%]
alpha_JET_EtaIntercalibration_NonClosure_posEta	±0.00 [0.00%]	±0.94 [0.10%]	±1.99 [0.20%]
alpha_Z_theory_fac	±0.00 [0.00%]	±5.68 [0.60%]	±8.94 [0.90%]
alpha_W_theory_renorm	±0.00 [0.00%]	±24.89 [2.6%]	±31.00 [3.1%]
alpha_JET_EtaIntercalibration_NonClosure_highE	±0.00 [0.00%]	±0.09 [0.01%]	±0.00 [0.00%]
alpha_W_theory_fac	±0.00 [0.00%]	±6.63 [0.70%]	±7.22 [0.72%]

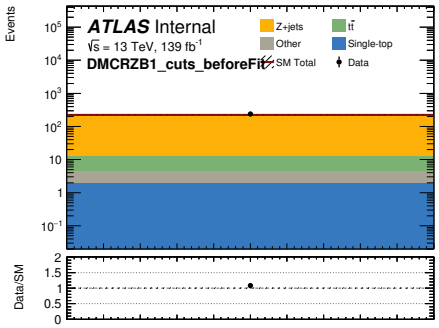
Table 7.2 Breakdown of the systematic uncertainties on background estimates for DM-CRZB1, DM-VRZB1 and DM-SRB1. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.



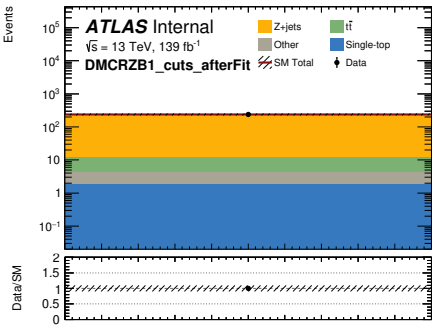
(a) NP pull plot for the background only fit for DM-SRB1

(b) NP correlations. NPs with an impact $< 1\%$ are removed from the plot.

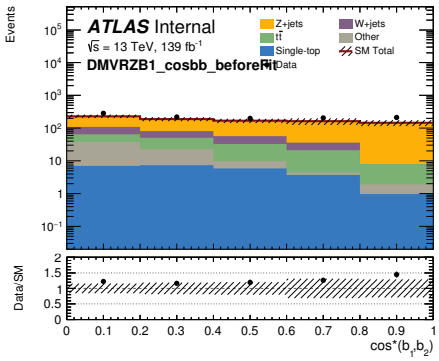
Fig. 7.1 Nuisance parameter plots for the stability of the background-only fit.



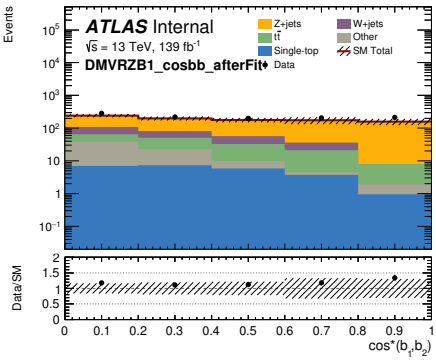
(a) DM-CRZB1, before fit



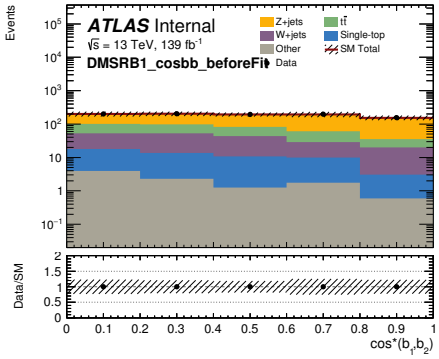
(b) DM-CRZB1, after fit



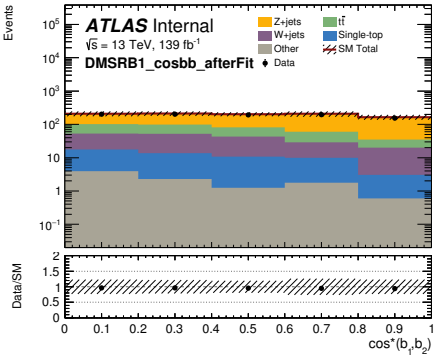
(c) DM-VRZB1, $\cos^*(b_1, b_2)$, before fit



(d) DM-VRZB1, $\cos^*(b_1, b_2)$, after fit



(e) DM-SRB1, $\cos^*(b_1, b_2)$, before fit

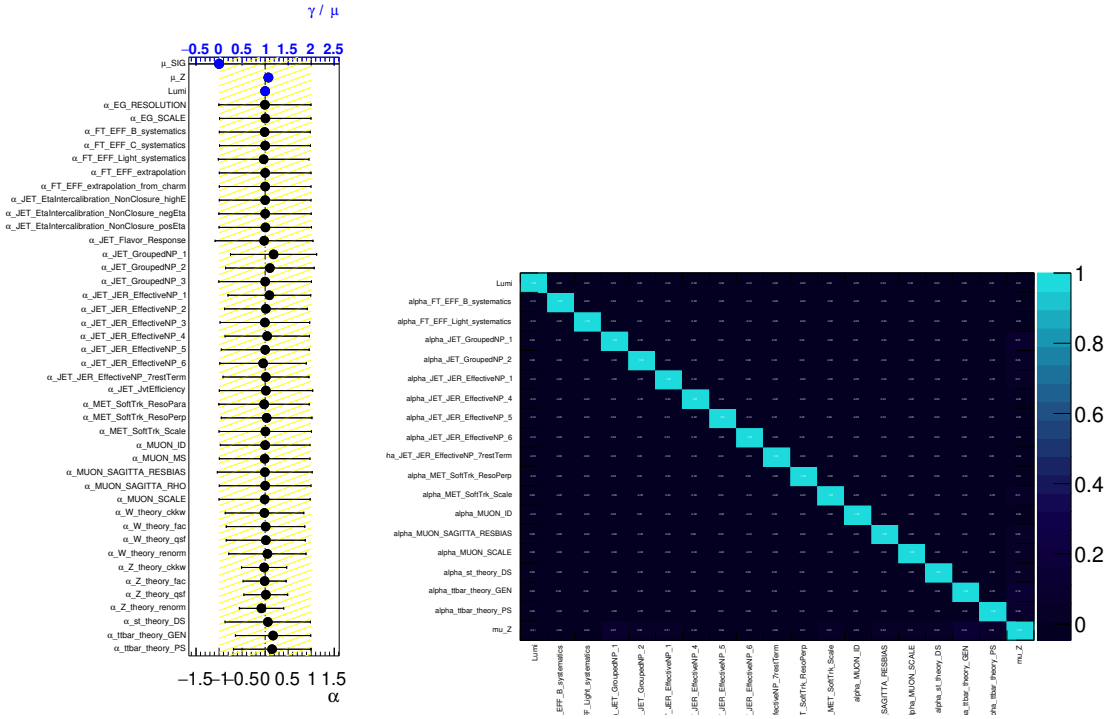


(f) DM-SRB1, $\cos^*(b_1, b_2)$, after fit

Fig. 7.2 CR, VR and SR plots pre-fit and post-fit. The hatched area corresponds to the combined systematic and statistical uncertainty on the MC backgrounds. The left column shows the pre-fit distributions and the right column shows the distributions post-fit. In the SR, data is blinded and set equal to the pre-fit SM total.

7.1.2 Exclusion fit results

The exclusion fit for DM-SRB1 is performed on five SR bins of the variable $\cos^*(b_1,b_2)$. Figure 7.3 shows that this fit is generally stable (i.e. the systematics are mostly run between 1 and -1 and are centered on 0) when a signal of $m_\phi = 20$ GeV is fitted to the signal strength parameter of interest, μ_{SIG}. Some constraining and pulling is observed however there are no pulls $> 1\sigma$. The plot shows that the signal strength parameter is consistent with 0 (i.e. there being no signal present) and that it can take any value within a large range of values that extend beyond the range of the axis. It must be noted that the results are blinded here and this is just a comparison of the pre-fit MC with the post-fit MC.



(a) NP pull plot for the exclusion fit for DM-SRB1.

(b) NP correlations. NPs with an impact of $< 1\%$ are removed from the plot.

Fig. 7.3 Nuisance parameter plots for the stability of the exclusion fit. The signal strength parameter, μ_{SIG}, has been included here.

The pre-fit and post-fit yields for the SR broken down into the five bins of $\cos^*(b_1,b_2)$ are shown in Table 7.3. The systematic impact in the SR in the exclusion fit are shown in Table 7.4.

	DMSRB1_cosbb	DMSRB1_cosbb_bin0	DMSRB1_cosbb_bin1	DMSRB1_cosbb_bin2	DMSRB1_cosbb_bin3	DMSRB1_cosbb_bin4
Observed events	950	201	203	194	196	154
Fitted bkg events	953.81 ± 53.44	201.0 ± 12.7	203.7 ± 11.2	194.4 ± 12.5	198.0 ± 13.5	156.8 ± 15.4
Fitted Z events	606.71 ± 51.01	106.2 ± 10.5	112.9 ± 12.2	118.2 ± 10.4	143.2 ± 15.4	126.2 ± 9.4
Fitted W events	138.33 ± 30.10	35.0 ± 9.2	37.5 ± 8.9	31.8 ± 6.8	18.0 ± 5.3	16.0 ± 5.3
Fitted ttbar events	155.13 ± 55.17	42.4 ± 14.1	39.8 ± 13.4	33.9 ± 10.8	27.5 ± 9.6	11.5 ± 9.3
Fitted st events	42.73 ± 33.58	13.2 ± 10.3	10.8 ± 8.0	8.8 ± 6.7	7.5 ± 7.4	2.3 ± 1.6
Fitted ttW events	0.07 ± 0.01	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Fitted ttZ events	0.39 ± 0.11	0.1 ± 0.0	0.0 ± 0.0	0.2 ± 0.1	0.0 ± 0.0	0.0 ± -0.0
Fitted ttH events	0.93 ± 0.17	0.3 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.2 ± 0.0	0.0 ± 0.0
Fitted diboson events	9.52 ± 1.68	3.8 ± 0.9	2.3 ± 0.3	1.2 ± 0.3	1.6 ± 1.0	0.6 ± 0.5
MC exp. SM events	956.62	202.66	204.55	195.46	197.58	156.37
MC exp. Z events	571.70	99.97	105.16	111.79	135.56	119.22
MC exp. W events	142.41	35.21	38.65	32.53	19.01	17.01
MC exp. ttbar events	179.45	48.31	45.79	38.68	31.50	15.17
MC exp. st events	45.50	13.83	11.43	9.54	8.22	2.46
MC exp. ttW events	0.07	0.04	0.01	0.02	0.00	0.00
MC exp. ttZ events	0.40	0.10	0.05	0.18	0.03	0.04
MC exp. ttH events	0.92	0.25	0.22	0.22	0.19	0.04
MC exp. diboson events	9.83	3.92	2.29	1.25	1.76	0.60

Table 7.3 DM-SRB1 exclusion fit yields. Fit results for the five $\cos^*(b_1, b_2)$ bins in the SR. Nominal MC expectations (i.e. pre-fit) are given for comparison. The errors shown are the statistical plus systematic uncertainties.

7.2 DM-SRB2 Fit Results

7.2.1 Background Only Fit Results

The structure of the fit for DM-SRB2, the region that is targeting high DM mediator masses (≥ 100 GeV), is identical to that of DM-SRB1. A single CR for the $Z+$jets background is used. A value for μ is obtained for the $Z+$jets background, μ^Z_{B1}, from DM-CRZB2, where DM-CRZB2 is as defined in Section 6.2. DM-VRZB2, defined in Section 6.3, is used to test the validity of this background estimation in a region that is kinematically similar to the SR. The normalisation factor obtained using DM-CRZB2 is

$$\mu^Z_{\mathrm{B2}} = 1.11 \pm 0.06 \tag{7.2}$$

Table 7.5 shows the pre-fit and post-fit yields for DM-CRZB2, DM-VRZB2 and DM-SRB2. The systematic breakdown in each region is given in Table 7.6.

The NP pulls are shown in Figure 7.4.

A series of pre-fit and post-fit distributions in the CR, VR and SR can be found in Figure 7.5.

Uncertainty of channel	DM-SRB1
Total background expectation	953.81
Total statistical ($\sqrt{N_{\text{exp}}}$)	±30.88
Total background systematic	±31.58 [3.31%]
alpha_ttbar_theory_PS	±51.34 [5.4%]
alpha_ttbar_theory_GEN	±48.87 [5.1%]
mu_Z	±40.51 [4.2%]
alpha_JET_GroupedNP_1	±35.15 [3.7%]
alpha_st_theory_DS	±33.83 [3.5%]
alpha_JET_GroupedNP_2	±28.50 [3.0%]
alpha_Z_theory_renorm	±23.98 [2.5%]
alpha_W_theory_renorm	±21.80 [2.3%]
alpha_JET_JER_EffectiveNP_6	±19.21 [2.0%]
alpha_JET_JER_EffectiveNP_2	±14.90 [1.6%]
alpha_FT_EFF_B_systematics	±11.33 [1.2%]
alpha_JET_JER_EffectiveNP_7restTerm	±11.25 [1.2%]
alpha_W_theory_qsf	±8.95 [0.94%]
alpha_W_theory_ckkw	±8.05 [0.84%]
alpha_FT_EFF_Light_systematics	±7.68 [0.80%]
alpha_JET_JER_EffectiveNP_5	±7.24 [0.76%]
alpha_FT_EFF_C_systematics	±6.53 [0.68%]
alpha_JET_JER_EffectiveNP_1	±6.51 [0.68%]
Lumi	±5.87 [0.62%]
alpha_Z_theory_ckkw	±5.74 [0.60%]
alpha_JET_JER_EffectiveNP_4	±5.73 [0.60%]
alpha_W_theory_fac	±5.43 [0.57%]
alpha_JET_JvtEfficiency	±5.19 [0.54%]
alpha_JET_JER_EffectiveNP_3	±5.02 [0.53%]
alpha_Z_theory_qsf	±4.03 [0.42%]
alpha_MET_SoftTrk_ResoPara	±2.80 [0.29%]
alpha_MET_SoftTrk_Scale	±2.76 [0.29%]
alpha_Z_theory_fac	±2.35 [0.25%]
alpha_JET_EtaIntercalibration_NonClosure_posEta	±2.00 [0.21%]
alpha_MUON_MS	±1.43 [0.15%]
alpha_MUON_SCALE	±1.14 [0.12%]
alpha_MUON_ID	±1.08 [0.11%]
alpha_JET_GroupedNP_3	±1.00 [0.11%]
alpha_EG_SCALE	±1.00 [0.10%]
alpha_JET_Flavor_Response	±0.84 [0.09%]
alpha_JET_EtaIntercalibration_NonClosure_negEta	±0.55 [0.06%]
alpha_EG_RESOLUTION	±0.38 [0.04%]
alpha_MET_SoftTrk_ResoPerp	±0.31 [0.03%]
alpha_MUON_SAGITTA_RESBIAS	±0.20 [0.02%]
alpha_MUON_SAGITTA_RHO	±0.00 [0.00%]
alpha_FT_EFF_extrapolation	±0.00 [0.00%]
mu_SIG	±0.00 [0.00%]
alpha_FT_EFF_extrapolation_from_charm	±0.00 [0.00%]
alpha_JET_EtaIntercalibration_NonClosure_highE	±0.00 [0.00%]

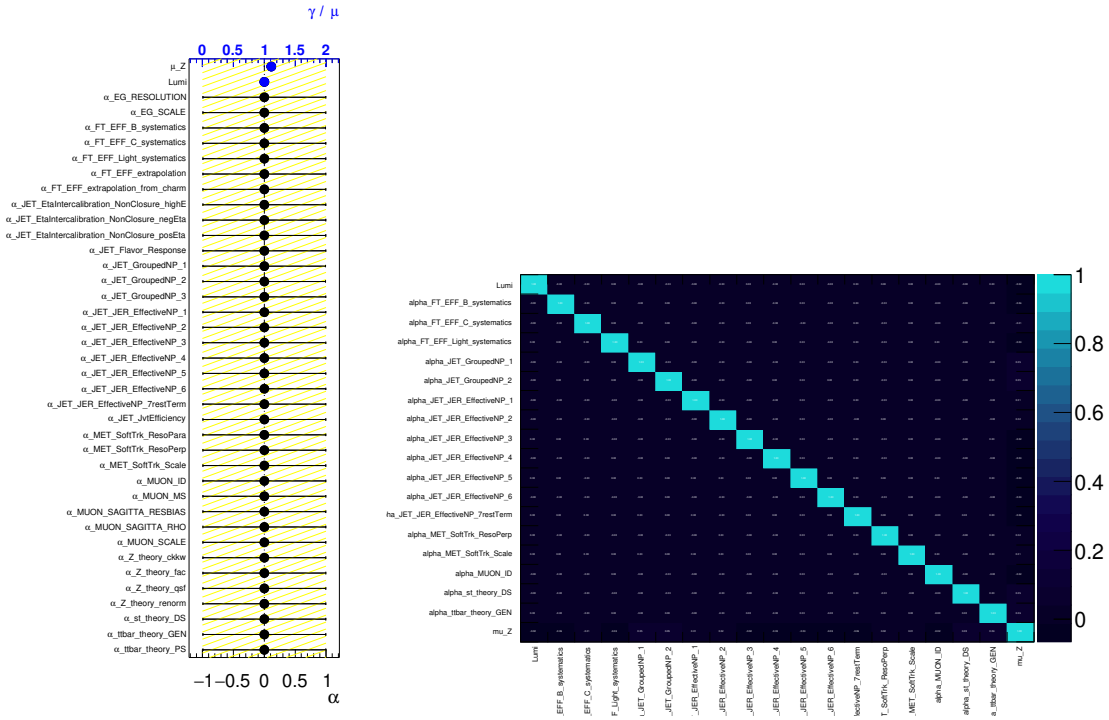
Table 7.4 Breakdown of the systematic uncertainties on background estimates for DM-SRB1. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.

	DM-CRZB2	DM-VRZB2	DM-SRB2
Observed events	433	1408	1623
Fitted bkg events	433.00 ± 20.80	1166.95 ± 150.74	1748.98 ± 225.42
Fitted Z events	407.26 ± 21.11	814.87 ± 142.36	1237.03 ± 179.44
Fitted W events	0.00 ± 0.00	174.18 ± 72.02	233.46 ± 81.93
Fitted ttbar events	7.99 ± 2.67	95.49 ± 37.37	158.91 ± 73.95
Fitted st events	2.88 ± 1.17	20.26 ± 13.24	45.05 ± 33.84
Fitted ttW events	0.02 ± 0.01	0.03 ± 0.02	0.13 ± 0.02
Fitted ttZ events	0.30 ± 0.06	0.12 ± 0.12	0.76 ± 0.23
Fitted ttH events	0.00 ± 0.00	0.45 ± 0.15	2.10 ± 0.32
Fitted diboson events	14.55 ± 1.18	61.55 ± 8.01	71.55 ± 6.52
MC exp. SM events	391.53	1083.95	1623.00
MC exp. Z events	365.79	731.88	1111.05
MC exp. W events	0.00	174.17	233.46
MC exp. ttbar events	7.99	95.49	158.90
MC exp. st events	2.88	20.26	45.05
MC exp. ttW events	0.02	0.03	0.13
MC exp. ttZ events	0.30	0.12	0.76
MC exp. ttH events	0.00	0.45	2.10
MC exp. diboson events	14.55	61.55	71.55

Table 7.5 Fit results for one bin DM-CRZB2, DM-VRZB2 and DM-SRB2. The results are obtained from the CRs using the background-only fit. Nominal MC expectations (i.e. pre-fit) are given for comparison. The errors shown are the statistical plus systematic uncertainties.

Uncertainty of channel	DMCRZB2	DMVRZB2	DMSRB2
Total background expectation	433.00	1166.95	1748.98
Total statistical ($\sqrt{N_{\text{exp}}}$)	±20.81	±34.16	±41.82
Total background systematic	±20.80 [4.80%]	±150.74 [12.92%]	±225.42 [12.89%]
mu_Z	±21.11 [4.9%]	±42.23 [3.6%]	±64.11 [3.7%]
alpha_ttbar_theory_GEN	±1.83 [0.42%]	±22.46 [1.9%]	±51.07 [2.9%]
alpha_FT_EFF_B_systematics	±1.36 [0.31%]	±24.98 [2.1%]	±33.82 [1.9%]
alpha_JET_JER_EffectiveNP_7restTerm	±1.16 [0.27%]	±25.00 [2.1%]	±27.61 [1.6%]
alpha_JET_GroupedNP_2	±1.11 [0.26%]	±35.11 [3.0%]	±20.79 [1.2%]
alpha_JET_GroupedNP_1	±1.11 [0.26%]	±27.00 [2.3%]	±30.32 [1.7%]
alpha_st_theory_DS	±0.98 [0.23%]	±12.83 [1.1%]	±33.09 [1.9%]
alpha_JET_JER_EffectiveNP_6	±0.73 [0.17%]	±4.68 [0.40%]	±39.81 [2.3%]
alpha_FT_EFF_Light_systematics	±0.59 [0.14%]	±11.75 [1.0%]	±19.53 [1.1%]
alpha_JET_JER_EffectiveNP_5	±0.56 [0.13%]	±2.43 [0.21%]	±38.64 [2.2%]
alpha_JET_JER_EffectiveNP_4	±0.53 [0.12%]	±6.54 [0.56%]	±34.45 [2.0%]
Lumi	±0.44 [0.10%]	±5.96 [0.51%]	±8.66 [0.50%]
alpha_MUON_ID	±0.43 [0.10%]	±0.60 [0.05%]	±3.03 [0.17%]
alpha_MET_SoftTrk_ResoPerp	±0.40 [0.09%]	±5.69 [0.49%]	±8.61 [0.49%]
alpha_JET_JER_EffectiveNP_3	±0.34 [0.08%]	±8.58 [0.74%]	±32.26 [1.8%]
alpha_JET_JER_EffectiveNP_2	±0.34 [0.08%]	±10.68 [0.92%]	±28.68 [1.6%]
alpha_JET_JER_EffectiveNP_1	±0.28 [0.07%]	±11.90 [1.0%]	±23.46 [1.3%]
alpha_FT_EFF_C_systematics	±0.24 [0.06%]	±11.80 [1.0%]	±16.91 [0.97%]
alpha_MET_SoftTrk_Scale	±0.23 [0.05%]	±14.31 [1.2%]	±0.47 [0.03%]
alpha_EG_RESOLUTION	±0.20 [0.05%]	±0.77 [0.07%]	±1.44 [0.08%]
alpha_EG_SCALE	±0.19 [0.04%]	±3.54 [0.30%]	±3.58 [0.20%]
alpha_MET_SoftTrk_ResoPara	±0.19 [0.04%]	±1.87 [0.16%]	±6.78 [0.39%]
alpha_MUON_SCALE	±0.13 [0.03%]	±1.36 [0.12%]	±2.01 [0.12%]
alpha_JET_Flavor_Response	±0.10 [0.02%]	±0.32 [0.03%]	±9.98 [0.57%]
alpha_MUON_SAGITTA_RESBIAS	±0.07 [0.02%]	±0.16 [0.01%]	±0.35 [0.02%]
alpha_MUON_MS	±0.07 [0.02%]	±0.83 [0.07%]	±3.52 [0.20%]
alpha_JET_GroupedNP_3	±0.07 [0.02%]	±7.29 [0.62%]	±9.17 [0.52%]
alpha_ttbar_theory_PS	±0.05 [0.01%]	±26.19 [2.2%]	±47.75 [2.7%]
alpha_JET_JvtEfficiency	±0.02 [0.00%]	±1.61 [0.14%]	±3.17 [0.18%]
alpha_JET_EtaIntercalibration_NonClosure_posEta	±0.01 [0.00%]	±0.24 [0.02%]	±0.30 [0.02%]
alpha_JET_EtaIntercalibration_NonClosure_negEta	±0.01 [0.00%]	±0.09 [0.01%]	±1.45 [0.08%]
alpha_MUON_SAGITTA_RHO	±0.00 [0.00%]	±0.15 [0.01%]	±0.03 [0.00%]
alpha_W_theory_ckkw	±0.00 [0.00%]	±13.36 [1.1%]	±18.17 [1.0%]
alpha_FT_EFF_extrapolation	±0.00 [0.00%]	±0.00 [0.00%]	±0.00 [0.00%]
alpha_W_theory_qsf	±0.00 [0.00%]	±22.42 [1.9%]	±19.23 [1.1%]
alpha_Z_theory_renorm	±0.00 [0.00%]	±107.26 [9.2%]	±152.92 [8.7%]
gamma_stat_DMVRZB2_cosbb_bin_1	±0.00 [0.00%]	±26.84 [2.3%]	±0.00 [0.00%]
gamma_stat_DMSRB2_cosbb_bin_4	±0.00 [0.00%]	±0.00 [0.00%]	±22.39 [1.3%]
alpha_FT_EFF_extrapolation_from_charm	±0.00 [0.00%]	±0.00 [0.00%]	±0.00 [0.00%]
alpha_Z_theory_ckkw	±0.00 [0.00%]	±15.74 [1.3%]	±39.00 [2.2%]
alpha_Z_theory_qsf	±0.00 [0.00%]	±29.43 [2.5%]	±35.51 [2.0%]
gamma_stat_DMVRZB2_cosbb_bin_4	±0.00 [0.00%]	±9.46 [0.81%]	±0.00 [0.00%]
alpha_Z_theory_fac	±0.00 [0.00%]	±5.54 [0.47%]	±19.27 [1.1%]
alpha_W_theory_renorm	±0.00 [0.00%]	±40.90 [3.5%]	±52.65 [3.0%]
alpha_JET_EtaIntercalibration_NonClosure_highE	±0.00 [0.00%]	±0.15 [0.01%]	±0.03 [0.00%]
alpha_W_theory_fac	±0.00 [0.00%]	±10.44 [0.89%]	±11.27 [0.64%]

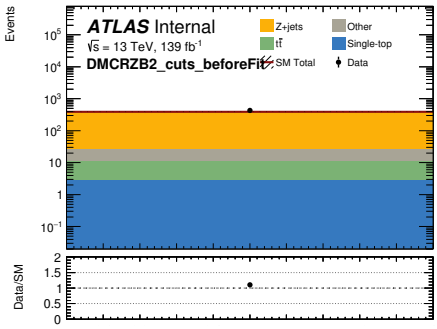
Table 7.6 Breakdown of the systematic uncertainties on background estimates in the DM-CRZB2, DM-VRZB2 and DM-SRB2. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.



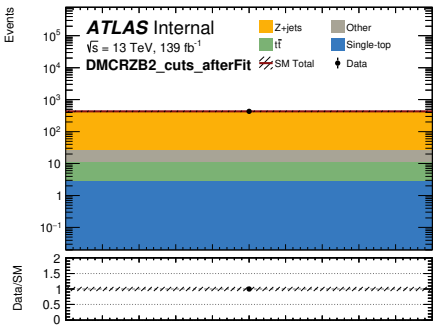
(a) NP pull plot for the background only fit for DM-SRB2.

(b) NP correlations. NPs with an impact of $< 1\%$ are removed from the plot.

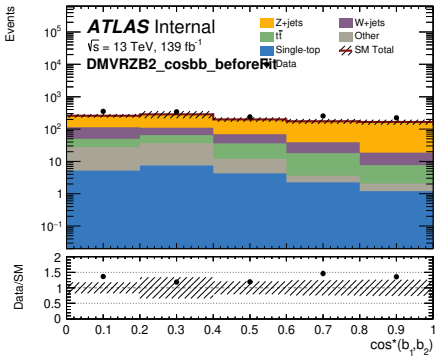
Fig. 7.4 Nuisance parameter plots for the stability of the background-only fit.



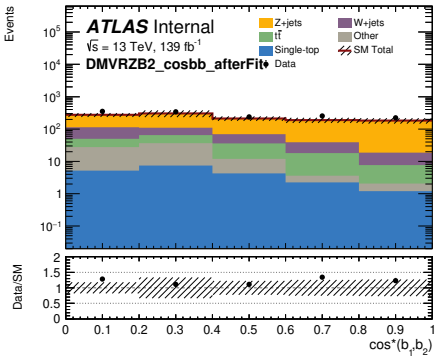
(a) DM-CRZB2, before fit



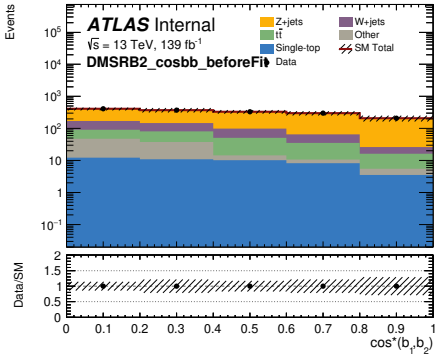
(b) DM-CRZB2, after fit



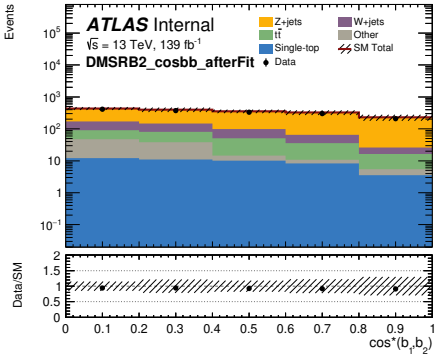
(c) DM-VRZB2, $\cos^*(b_1, b_2)$, before fit



(d) DM-VRZB2, $\cos^*(b_1, b_2)$, after fit



(e) DM-SRB2, $\cos^*(b_1, b_2)$, before fit

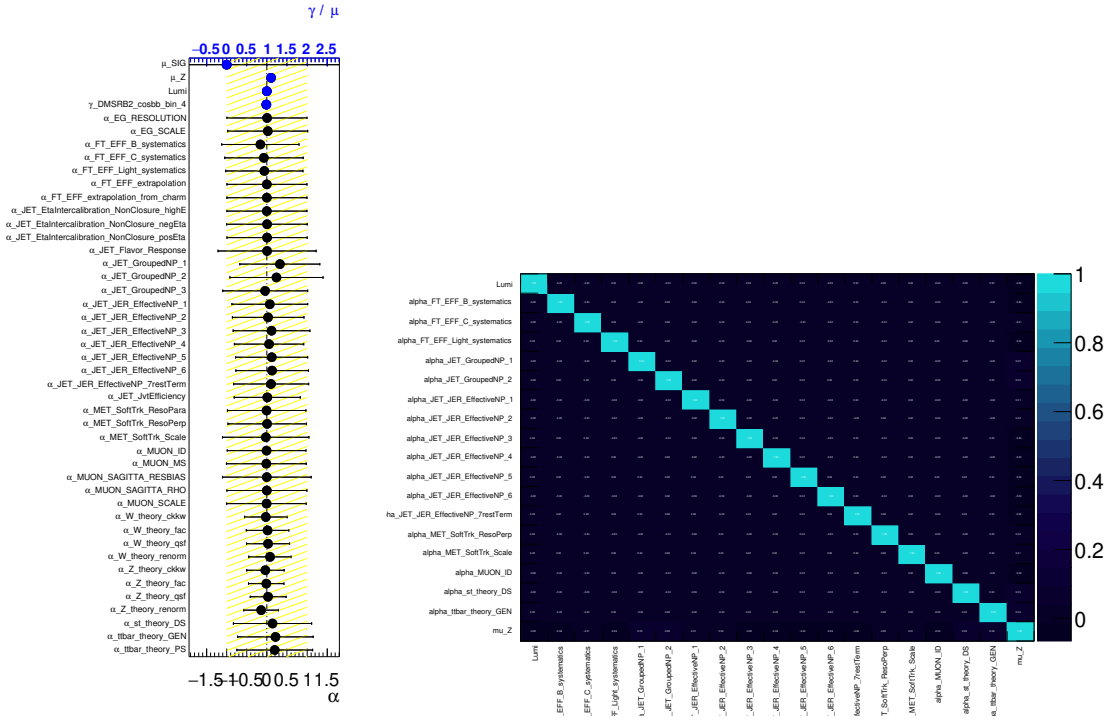


(f) DM-SRB2, $\cos^*(b_1, b_2)$, after fit

Fig. 7.5 CR, VR and SR plots pre-fit and post-fit. The hatched area corresponds to the combined systematic and statistical uncertainty on the MC backgrounds. The left column shows the pre-fit distributions and the right column shows the distributions post-fit. In the SR, data is blinded and set equal to the pre-fit SM total.

7.2.2 Exclusion fit results

The exclusion fit for DM-SRB2 is performed on five SR bins of the variable $\cos^*(b_1,b_2)$. Figure 7.6 shows that this fit is stable when a signal of $m_\phi = 20$ GeV is fitted to the signal strength parameter of interest, μ_{SIG}.



(a) NP pull plot for the exclusion fit for DM-SRB2.

(b) NP correlations. NPs with an impact of $< 1\%$ are removed from the plot.

Fig. 7.6 Nuisance parameter plots for the stability of the exclusion fit. The signal strength parameter, μ_{SIG}, has been included here.

The pre-fit and post-fit yields for the SR, broken down into the five bins of $\cos^*(b_1,b_2)$, are shown in Table 7.7. The impact of each systematic in the SR in the exclusion fit is shown in Table 7.8.

	DM-SRB2
Observed events	1623
Fitted bkg events	1630.96 ± 69.73
Fitted Z events	1187.54 ± 63.91
Fitted W events	208.30 ± 41.06
Fitted ttbar events	126.55 ± 57.69
Fitted st events	36.85 ± 31.66
Fitted ttW events	0.12 ± 0.02
Fitted ttZ events	0.68 ± 0.15
Fitted ttH events	1.99 ± 0.27
Fitted diboson events	68.93 ± 5.12
MC exp. SM events	1628.02
MC exp. Z events	1111.05
MC exp. W events	233.46
MC exp. ttbar events	158.90
MC exp. st events	45.05
MC exp. ttW events	0.13
MC exp. ttZ events	0.76
MC exp. ttH events	2.10
MC exp. diboson events	71.55

Table 7.7 DM-SRB2 exclusion fit yields. Fit results for the five $\cos^*(b_1, b_2)$ bins in the SR. Nominal MC expectations (i.e. pre-fit) are given for comparison. The errors shown are the statistical plus systematic uncertainties.

Uncertainty of channel	DM-SRB2
Total background expectation	1630.96
Total statistical ($\sqrt{N_{\text{exp}}}$)	±40.39
Total background systematic	±51.73 [3.17%]
mu_Z	±59.36 [3.6%]
alpha_ttbar_theory_GEN	±48.33 [3.0%]
alpha_JET_GroupedNP_1	±46.30 [2.8%]
alpha_ttbar_theory_PS	±45.25 [2.8%]
alpha_JET_JER_EffectiveNP_6	±35.77 [2.2%]
alpha_JET_JER_EffectiveNP_5	±34.44 [2.1%]
alpha_JET_GroupedNP_2	±33.42 [2.0%]
alpha_FT_EFF_B_systematics	±32.40 [2.0%]
alpha_st_theory_DS	±32.19 [2.0%]
alpha_JET_JER_EffectiveNP_3	±30.43 [1.9%]
alpha_Z_theory_renorm	±29.49 [1.8%]
alpha_JET_JER_EffectiveNP_4	±29.39 [1.8%]
alpha_JET_JER_EffectiveNP_7restTerm	±25.44 [1.6%]
alpha_JET_JER_EffectiveNP_2	±25.40 [1.6%]
alpha_JET_JER_EffectiveNP_1	±22.03 [1.4%]
alpha_FT_EFF_Light_systematics	±19.20 [1.2%]
alpha_FT_EFF_C_systematics	±16.58 [1.0%]
gamma_stat_DMSRB2_cosbb_bin_4	±15.66 [0.96%]
alpha_W_theory_renorm	±14.50 [0.89%]
alpha_JET_Flavor_Response	±12.06 [0.74%]
alpha_JET_GroupedNP_3	±10.28 [0.63%]
alpha_Z_theory_ckkw	±8.53 [0.52%]
alpha_MET_SoftTrk_ResoPerp	±8.13 [0.50%]
Lumi	±7.50 [0.46%]
alpha_Z_theory_qsf	±6.78 [0.42%]
alpha_MET_SoftTrk_ResoPara	±6.54 [0.40%]
alpha_W_theory_qsf	±5.61 [0.34%]
alpha_W_theory_ckkw	±5.25 [0.32%]
alpha_Z_theory_fac	±3.52 [0.22%]
alpha_EG_SCALE	±3.45 [0.21%]
alpha_MUON_MS	±3.38 [0.21%]
alpha_W_theory_fac	±3.25 [0.20%]
alpha_MUON_ID	±2.86 [0.18%]
alpha_JET_JvtEfficiency	±2.38 [0.15%]
alpha_MUON_SCALE	±1.93 [0.12%]
alpha_JET_EtaIntercalibration_NonClosure_negEta	±1.42 [0.09%]
alpha_EG_RESOLUTION	±1.41 [0.09%]
alpha_MET_SoftTrk_Scale	±0.68 [0.04%]
alpha_MUON_SAGITTA_RESBIAS	±0.41 [0.02%]
alpha_JET_EtaIntercalibration_NonClosure_posEta	±0.33 [0.02%]
alpha_JET_EtaIntercalibration_NonClosure_highE	±0.03 [0.00%]
alpha_MUON_SAGITTA_RHO	±0.03 [0.00%]
alpha_FT_EFF_extrapolation	±0.00 [0.00%]
mu_SIG	±0.00 [0.00%]
alpha_FT_EFF_extrapolation_from_charm	±0.00 [0.00%]

Table 7.8 Breakdown of the systematic uncertainties on background estimates for DM-SRB2. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.

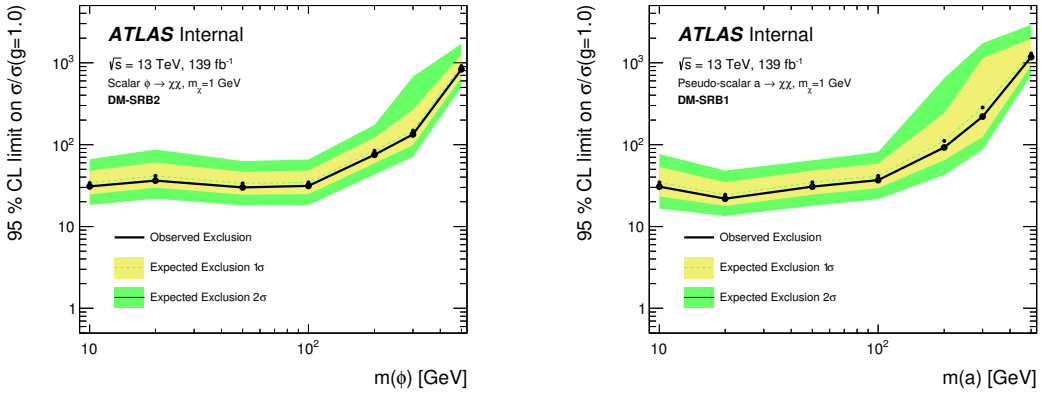
7.3 Expected Exclusion Results

The expected exclusion is calculated using the CLs method [94], which is detailed here. The probability that the data is consistent with the combined signal+background hypothesis is defined as $p_{\mathrm{s+b}}$ and the probability that the data is consistent with the background only hypothesis is p_{b}. The CLs value is defined as

$$\mathrm{CLs} = \frac{p_{\mathrm{s+b}}}{1 - p_{\mathrm{b}}}. \tag{7.3}$$

By convention, a model is excluded for CLs < 0.05. The y-axis of the figures presented here show the factor that the nominal cross section ($g = 1$) must be multiplied by to obtain a CLs of 0.05. The mass of the DM mediator is on the x-axis. Figures here show only the expected exclusion after the fit. As these results are blinded, the pre-fit MC is used as a place-holder for the data. This is shown as a solid black line in the figures. The yellow band corresponds to $\pm 1\sigma$ uncertainty and the green band corresponds to $\pm 2\sigma$ uncertainty on this exclusion. Everything above the dashed line is expected to be excluded.

Figure 7.7 shows the expected cross section exclusion at the 95% CL for scalar and pseudo-scalar DM mediators for DM-SRB1.



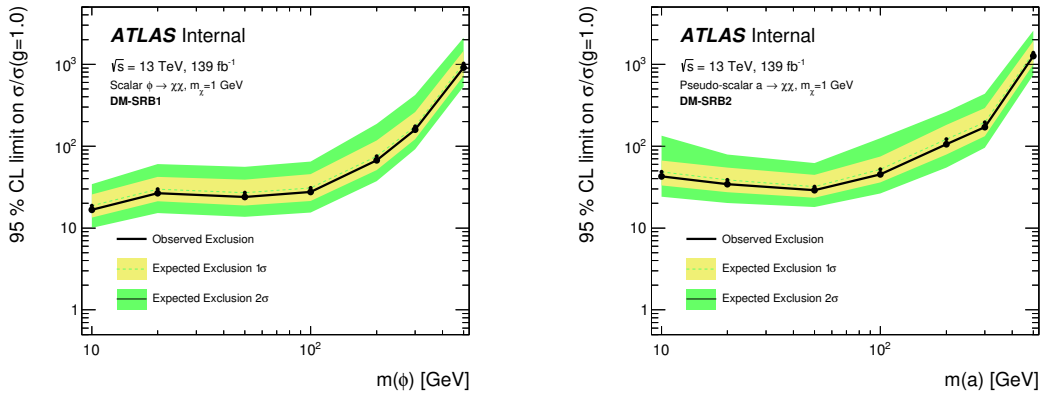
(a) Expected exclusion for scalar DM mediators

(b) Expected exclusion for pseudo-scalar DM mediators

Fig. 7.7 Expected cross section limits as a function of the nominal cross section value ($g = 1$) for (a) scalar DM mediatos and (b) pseudo-scalar DM mediators for DM-SRB1. Rather than showing the observed limit, the solid black line here shows the pre-fit expected exclusion.

Figure 7.8 shows the expected cross section exclusion for scalar and pseudo-scalar DM mediators for DM-SRB2.

Figure 7.9 shows the expected cross section exclusion for scalar and pseudo-scalar DM mediators for DM-SRB1 and DM-SRB2 combined. These regions have not been designed



(a) Expected exclusion for scalar DM mediators

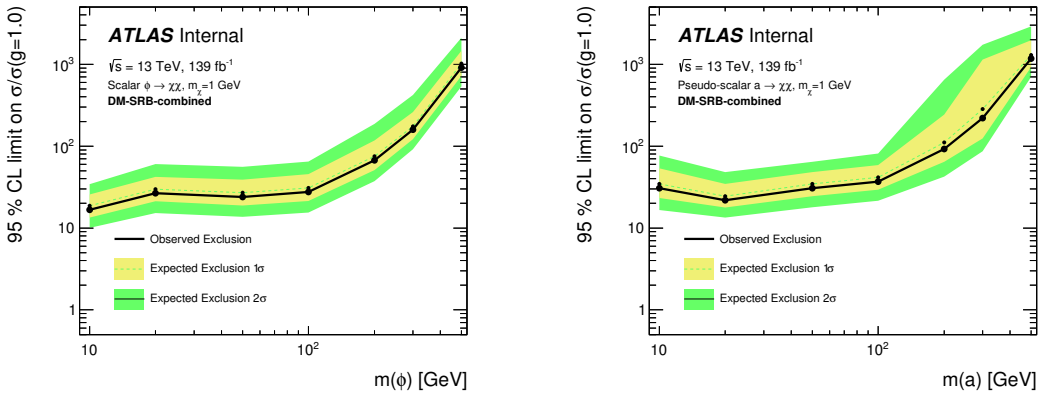
(b) Expected exclusion for pseudo-scalar DM mediators

Fig. 7.8 Expected cross section limits as a function of the nominal cross section value ($g = 1$) for (a) scalar DM mediators and (b) pseudo-scalar DM mediators for DM-SRB2. Rather than showing the observed limit, the solid black line here shows the pre-fit expected exclusion.

orthogonal to each other and therefore this is not a statistical combination. This plot is made by simply using the best expected limit from either SR. Generally this is DM-SRB1 for mediator masses $\leq$ 100 GeV and DM-SRB2 for mediator masses $>$ 100 GeV.

7.4 Outlook and Conclusions

This analysis represents, to date, the best expected lower limits on the production cross section of s-channel DM mediators produced in association with *b*-quarks. This limit is between 20 and 30 times the nominal cross section for scalar mediator masses between 10 and 100 GeV and between 30 and 40 times the nominal cross section for pseudo-scalar mediator masses in the same range. This assumes a dark matter mass of 1 GeV and a unitary coupling. This marks a great improvement on the previous observed limits set by ATLAS of 300 times the nominal cross section for mediator masses between 10 and 50 GeV [95]. This can partly be be attributed to the increase in luminosity of a factor four. Assuming only statistical uncertainties, this should correspond to an increase in sensitivity of the analysis of a factor two. A factor between five and ten is achieved here. This remarkable increase can be attributed to a number of factors. There have been ATLAS-wide developments of object definitions that have contributed greatly to this. It is not possible to mention every collaboration-wide improvement that has led to this improved result, but a particular mention should be made to the implementation of PFlow jets and DL1r *b*-tagging. On the level of this analysis, the use of BDTs in the SR definitions allow for a greater sensitivity to a signal



(a) Expected exclusion for scalar DM mediators

(b) Expected exclusion for pseudo-scalar DM mediators

Fig. 7.9 Expected cross section limits as a function of the nominal cross section value ($g = 1$) for (a) scalar DM mediators and (b) pseudo-scalar DM mediators for DM-SRB2. Rather than showing the observed limit, the solid black line here shows the pre-fit expected exclusion.

process which is very difficult to separate from its backgrounds. Complimentary to this, making use of the correlation between E_T^{miss} and $p_T(j_1)$ meant that more signal statistics could be retained before defining the signal region.

There are further improvements that could have been made to potentially increase the sensitivity of this analysis. For example, the detector signature of $W+$jets and $Z+$jets is very similar for all hadronic decay modes. Combining these two samples for the training of the BDT would have increased the size of the sample the BDT was trained on and therefore might have improved the discriminatory power of the BDT. In addition to these kinds of improvements to the analysis strategy, it would also be interesting to go back and reassess the theoretical context of this model. For example, one can ask how well the simplified model approach is working. There are also potential extensions to this analysis that could be explored in the near future. For example the SRs defined in this thesis are expected to be highly sensitive to DM models with two Higgs doublets and a pseudo-scalar mediator [96] in some regions of the parameter space of this model.

Novel analysis techniques as well as collaboration-wide developments have led to the best ever expected limit for the production cross section of s-channel DM mediators produced in association with b-quarks, hugely outperforming the expected improvement due to the increase in available statistics from the full Run 2 data set. The ability to perform an analysis like this is only possible because of the concerted efforts thousands of physicists, engineers and technicians who have worked on the ATLAS experiment over several decades and the

centuries of hard work in High Energy Physics that preceded this. This thesis represents an exploration into a small corner of the vast and rich landscape of High Energy Physics.

References

[1] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, "The Standard model of particle physics," *Rev. Mod. Phys.*, vol. 71, pp. S96–S111, 1999.

[2] M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev.*, vol. D98, no. 3, p. 030001, 2018.

[3] M. Aker *et al.*, "An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN," *Phys. Rev. Lett.*, vol. 123, no. 22, p. 221802, 2019.

[4] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.*, vol. B716, pp. 1–29, 2012.

[5] S. Chatrchyan *et al.*, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," *Phys. Lett. B*, vol. 716, pp. 30–61, 2012.

[6] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964. [,157(1964)].

[7] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Lett.*, vol. 12, pp. 132–133, 1964.

[8] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964. [,160(1964)].

[9] S. Chatrchyan *et al.*, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," *Phys. Lett.*, vol. B716, pp. 30–61, 2012.

[10] A. Collaboration, "Standard model production cross section measurements," 2019.

[11] "Popular information." https://www.nobelprize.org/prizes/physics/2004/popular-information/. Accessed: 2020-01-23.

[12] A. D. Sakharov, "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe," *Pisma Zh. Eksp. Teor. Fiz.*, vol. 5, pp. 32–35, 1967. [Usp. Fiz. Nauk161,no.5,61(1991)].

[13] S. Pakvasa, "CP violation in the neutrino sector," *Journal of Physics: Conference Series*, vol. 556, p. 012060, nov 2014.

[14] P. Schmidt-Wellenburg, "The quest to find an electric dipole moment of the neutron," *Asian Journal of Physics*, 07 2016.

[15] A. Drlica-Wagner *et al.*, "Probing the Fundamental Nature of Dark Matter with the Large Synoptic Survey Telescope," 2019.

[16] V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, 2005.

[17] L. Guzzo, J. Bel, D. Bianchi, C. Carbone, B. R. Granett, A. J. Hawken, F. G. Mohammad, A. Pezzotta, S. Rota, and M. Zennaro, "Measuring the universe with galaxy redshift surveys," in *Toward a Science Campus in Milan* (P. F. Bortignon, G. Lodato, E. Meroni, M. G. Paris, L. Perini, and A. Vicini, eds.), (Cham), pp. 1–16, Springer International Publishing, 2018.

[18] W. L. Freedman *et al.*, "The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch," 2019.

[19] M. Tegmark *et al.*, "Cosmological parameters from SDSS and WMAP," *Phys. Rev. D*, vol. 69, p. 103501, 2004.

[20] N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," 2018.

[21] S. McGaugh, "A Precise Milky Way Rotation Curve Model for an Accurate Galactocentric Distance," 2018.

[22] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, "A direct empirical proof of the existence of dark matter," *Astrophys. J.*, vol. 648, pp. L109–L113, 2006.

[23] D. Paraficz, J. P. Kneib, J. Richard, A. Morandi, M. Limousin, E. Jullo, and J. Martinez, "The Bullet cluster at its best: weighing stars, gas, and dark matter," *Astron. Astrophys.*, vol. 594, p. A121, 2016.

[24] J. R. Brownstein and J. W. Moffat, "The Bullet Cluster 1E0657-558 evidence shows Modified Gravity in the absence of Dark Matter," *Mon. Not. Roy. Astron. Soc.*, vol. 382, pp. 29–47, 2007.

[25] A. Robertson, R. Massey, and V. Eke, "What does the Bullet Cluster tell us about self-interacting dark matter?," *Mon. Not. Roy. Astron. Soc.*, vol. 465, no. 1, pp. 569–587, 2017.

[26] "The matter of the bullet cluster." https://apod.nasa.gov/apod/ap170115.html. Accessed: 2019-08-19.

[27] S. Sarkar, "Measuring the baryon content of the universe," in *Proceedings, XIIIth Rencontres de Blois on Frontiers of the Universe: Blois, France, June 17-23, 2001*, pp. 53–63, 2004.

[28] M. Bradac, D. Clowe, A. H. Gonzalez, P. Marshall, W. Forman, C. Jones, M. Markevitch, S. Randall, T. Schrabback, and D. Zaritsky, "Strong and weak lensing united. 3. Measuring the mass distribution of the merging galaxy cluster 1E0657-56," *Astrophys. J.*, vol. 652, pp. 937–947, 2006.

[29] A. Burrows, M. T. Ressell, and M. S. Turner, "Axions and SN1987A: Axion trapping," *Phys. Rev.*, vol. D42, pp. 3297–3309, 1990.

[30] R. D. Peccei, "The Strong CP problem and axions," *Lect. Notes Phys.*, vol. 741, pp. 3–17, 2008. [,3(2006)].

[31] P. Jetzer, "Gravitational microlensing research," in *Recent developments in theoretical and experimental general relativity, gravitation, and relativistic field theories. Proceedings, 8th Marcel Grossmann meeting, MG8, Jerusalem, Israel, June 22-27, 1997. Pts. A, B*, pp. 1543–1560, 1997.

[32] R. Contino, A. Falkowski, F. Goertz, C. Grojean, and F. Riva, "On the Validity of the Effective Field Theory Approach to SM Precision Tests," *JHEP*, vol. 07, p. 144, 2016.

[33] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, "Minimal flavor violation: An Effective field theory approach," *Nucl. Phys.*, vol. B645, pp. 155–187, 2002.

[34] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN." Dec 2008.

[35] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003. 437 p, 2008. Also published by CERN Geneva in 2010.

[36] A. L. Rosa, "The ATLAS insertable b-layer: from construction to operation," *Journal of Instrumentation*, vol. 11, pp. C12036–C12036, dec 2016.

[37] J. J. Peña, "Alignment of the ATLAS inner detector upgraded for the LHC run II," *Journal of Physics: Conference Series*, vol. 664, p. 072025, dec 2015.

[38] A. Boldyrev *et al.*, "The ATLAS transition radiation tracker," *Instrum. Exp. Tech.*, vol. 55, pp. 323–334, 2012.

[39] A. Ball, "Technical challenges of the large hadron collider experiments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, jan 2015.

[40] The ATLAS Collaboration, "Luminosity public results run2," 2019. (accessed 2019-10-11).

[41] A. R. Martínez, "The Run-2 ATLAS Trigger System," *J. Phys. Conf. Ser.*, vol. 762, no. 1, p. 012003, 2016.

[42] Y. Nakahama, "The ATLAS trigger system: Ready for run-2," *Journal of Physics: Conference Series*, vol. 664, p. 082037, dec 2015.

[43] R. Achenbach *et al.*, "The ATLAS level-1 calorimeter trigger," *JINST*, vol. 3, p. P03001, 2008.

[44] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, "Calorimeter Clustering Algorithms: Description and Performance," Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[45] B. Wiik, "HERA," *Conf. Proc. C*, vol. 8405141, pp. 23–42, 1984.

[46] R. R. Wilson, "The Tevatron," *Phys. Today*, vol. 30N10, pp. 23–30, 1977.

[47] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.*, vol. B126, pp. 298–318, 1977.

[48] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972. [Yad. Fiz.15,781(1972)].

[49] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[50] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, "QCD matrix elements + parton showers," *JHEP*, vol. 11, p. 063, 2001.

[51] A. Annovi, I. Vivarelli, and G. Volpi, "Study on $Z \to b\bar{b}$ measurement," 2004.

[52] B. Andersson, S. Mohanty, and F. Soderberg, "Recent developments in the Lund model," in *36th Annual Winter School on Nuclear and Particle Physics (PINP 2002) and 8th St. Petersburg School on Theoretical Physics St. Petersburg, Russia, February 25-March 3, 2002*, 2002.

[53] K. Odagiri, "Local charge compensation from color preconfinement as a key to the dynamics of hadronization," *JHEP*, vol. 07, p. 022, 2003.

[54] Z. Marshall, "Simulation of Pile-up in the ATLAS Experiment," *J. Phys. Conf. Ser.*, vol. 513, p. 022024, 2014.

[55] S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003.

[56] W. Lukas, "Fast Simulation for ATLAS: Atlfast-II and ISF," *J. Phys. Conf. Ser.*, vol. 396, p. 022031, 2012.

[57] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP*, vol. 06, p. 128, 2011.

[58] T. Sjostrand, "PYTHIA 8 Status Report," in *Proceedings, HERA and the LHC Workshop Series on the implications of HERA for LHC physics: 2006-2008*, pp. 726–732, 2008.

[59] R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP*, vol. 04, p. 040, 2015.

[60] Y. Afik, F. Maltoni, K. Mawatari, P. Pani, G. Polesello, Y. Rozen, and M. Zaro, "DM+$b\bar{b}$ simulations with DMSimp: an update," in *Dark Matter at the LHC 2018: Experimental and theoretical workshop (DM@LHC 2018) Heidelberg, Germany, April 3-6, 2018*, 2018.

[61] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, "FeynRules 2.0 - A complete toolbox for tree-level phenomenology," *Comput. Phys. Commun.*, vol. 185, pp. 2250–2300, 2014.

[62] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," *JHEP*, vol. 02, p. 007, 2009.

[63] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO," *Phys. Rev. Lett.*, vol. 103, p. 082001, 2009.

[64] C. Zhang, "NLO predictions for effective field theory with MadGraph5 aMC@NLO," *PoS*, vol. RADCOR2015, p. 101, 2016.

[65] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[66] "ATLAS Pythia 8 tunes to 7 TeV datas," Tech. Rep. ATL-PHYS-PUB-2014-021, The ATLAS Collaboration, Geneva, Nov 2014.

[67] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[68] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 06, p. 043, 2010.

[69] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$," *Phys. Rev. Lett.*, vol. 110, p. 252004, 2013.

[70] M. Czakon and A. Mitov, "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction," *JHEP*, vol. 01, p. 080, 2013.

[71] M. Czakon and A. Mitov, "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels," *JHEP*, vol. 12, p. 054, 2012.

[72] P. Bärnreuther, M. Czakon, and A. Mitov, "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$," *Phys. Rev. Lett.*, vol. 109, p. 132001, 2012.

[73] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation," *Phys. Lett.*, vol. B710, pp. 612–622, 2012.

[74] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders," *Comput. Phys. Commun.*, vol. 185, p. 2930, 2014.

[75] N. Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production," *Phys. Rev.*, vol. D83, p. 091503, 2011.

[76] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-," *Phys. Rev.*, vol. D82, p. 054018, 2010.

[77] N. Kidonakis, "NNLL resummation for s-channel single top quark production," *Phys. Rev.*, vol. D81, p. 054028, 2010.

[78] The ATLAS Collaboration, "Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC," *Eur. Phys. J.*, vol. C76, no. 12, p. 653, 2016.

[79] The ATLAS Collaboration, "Jet reconstruction and performance using particle flow with the ATLAS Detector," *Eur. Phys. J.*, vol. C77, no. 7, p. 466, 2017.

[80] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[81] "Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS," Tech. Rep. ATLAS-CONF-2015-037, The ATLAS Collaboration, Geneva, Aug 2015.

[82] M. Lanfermann, "Deep Learning in Flavour Tagging at the ATLAS experiment," *PoS*, vol. EPS-HEP2017, p. 764, 2018.

[83] J. Geisen, *Calibration of the ATLAS B-tagger and the search for the $t\bar{t}H(H \to b\bar{b})$ process at $\sqrt{s}$ = 13 TeV with the ATLAS experiment at the LHC*. PhD thesis, Gottingen U., 2019.

[84] "Electron identification measurements in ATLAS using $\sqrt{s}$ = 13 TeV data with 50 ns bunch spacing," Tech. Rep. ATL-PHYS-PUB-2015-041, The ATLAS Collaboration, Geneva, Sep 2015.

[85] C. Di Donato, "Muon identification and performance in the ATLAS experiment," *PoS*, vol. BEAUTY2018, p. 068, 2018.

[86] The ATLAS Collaboration, "Overlap Removal Details," 2019. (accessed 2020-02-13).

[87] The ATLAS Collaboration, "Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector," *Eur. Phys. J.*, vol. C77, no. 9, p. 580, 2017. [Erratum: Eur. Phys. J.C77,no.10,712(2017)].

[88] "Object-based missing transverse momentum significance in the ATLAS detector," Tech. Rep. ATLAS-CONF-2018-038, CERN, Geneva, Jul 2018.

[89] D. R. Tovey, "On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders," *JHEP*, vol. 04, p. 034, 2008.

[90] U. Haisch, P. Pani, and G. Polesello, "Determining the CP nature of spin-0 mediators in associated production of dark matter and $t\bar{t}$ pairs," *JHEP*, vol. 02, p. 131, 2017.

[91] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[92] Y. Coadou, "Boosted Decision Trees and Applications," *EPJ Web Conf.*, vol. 55, p. 02004, 2013.

[93] J. M. Lorenz, M. Baak, G. J. Besjes, D. Côté, A. Koutsman, and D. Short, "HistFitter - A flexible framework for statistical data analysis," *J. Phys. Conf. Ser.*, vol. 608, p. 012049, 2015.

[94] A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys. G*, vol. 28, pp. 2693–2704, 2002.

[95] M. Aaboud *et al.*, "Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector," *Eur. Phys. J.*, vol. C78, no. 1, p. 18, 2018.

[96] M. Bauer, U. Haisch, and F. Kahlhoefer, "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators," *JHEP*, vol. 05, p. 138, 2017.

[97] The ATLAS Collaboration, "PmgSystematicUncertaintyRecipes," 2004. [Online; accessed 22-July-2019].

Appendix A

DM BDT Validation

A.1 Data MC comparisons

In this Section the data MC comparisons are shown for all variables used in all DM SRs, CRs and VRs. Also shown are the BDT discriminant distributions in all regions.

A.2 BDT profile plots

The dependence of each kinematic variable used for the BDT trainings is shown for each BDT bin for DM-pre-2b-2j-0l-lowMET. The total MC background is compared to data and generally there is good agreement.

Figures A.10, A.11 and A.12 show shape comparisons for some important kinematic variables. There are strong correlations between these variables and the BDT output scores and so the ratios of the disributions in these regions is not expected to be flat.

A.3 n-1 BDT plots

So-called "n-1" plots show the distribution of a particular variable with a selectionapplied, but the selection on the plotted variable has been removed from the selection. Figures A.13 and A.14 show the n-1 distributions for DM BDTs trained on low mediator masses and high mediator masses respectively. The low mass BDTs are shown in DM-SRB1 and the high mass BDTs are shown in DM-SRB2.

Figures A.15 and A.16 show the BDT distributions in the CRs without the cut on the relevant $Z+$jetsBDT discriminant.

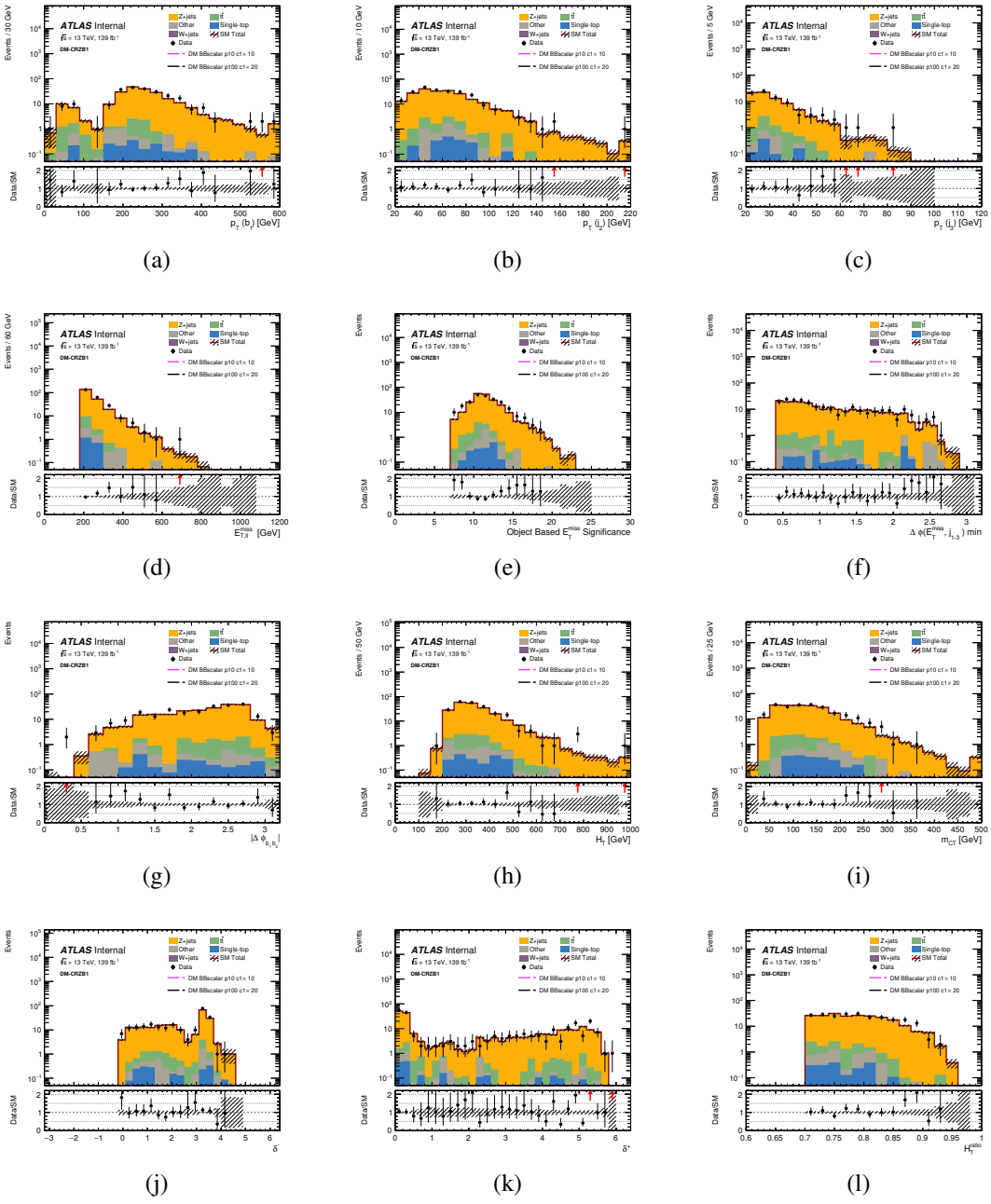


Fig. A.1 Data MC comparisons for all variables used in BDT training for DM-CRZB1.

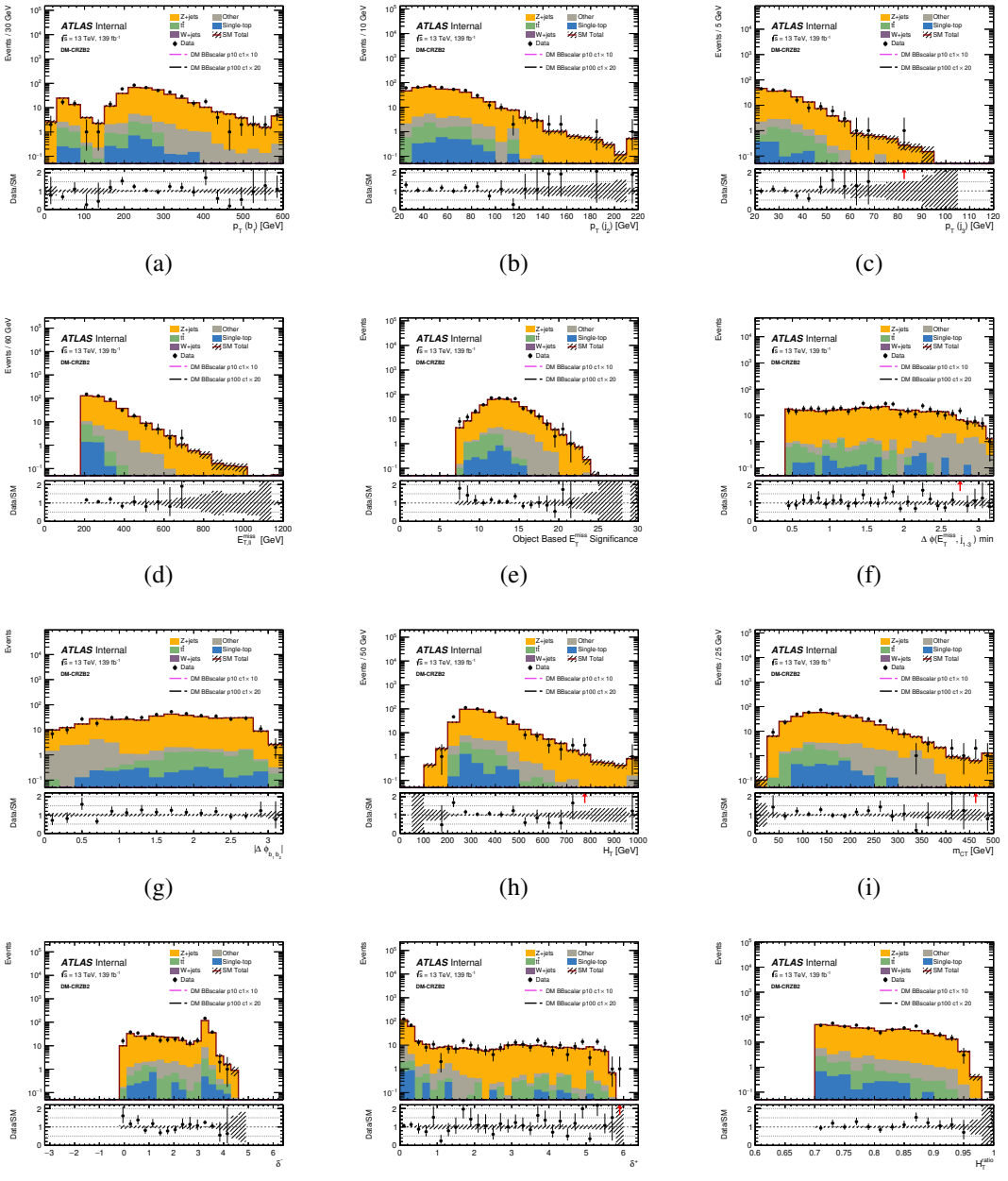


Fig. A.2 Data MC comparisons for all variables used in BDT training for DM-CRZB2.

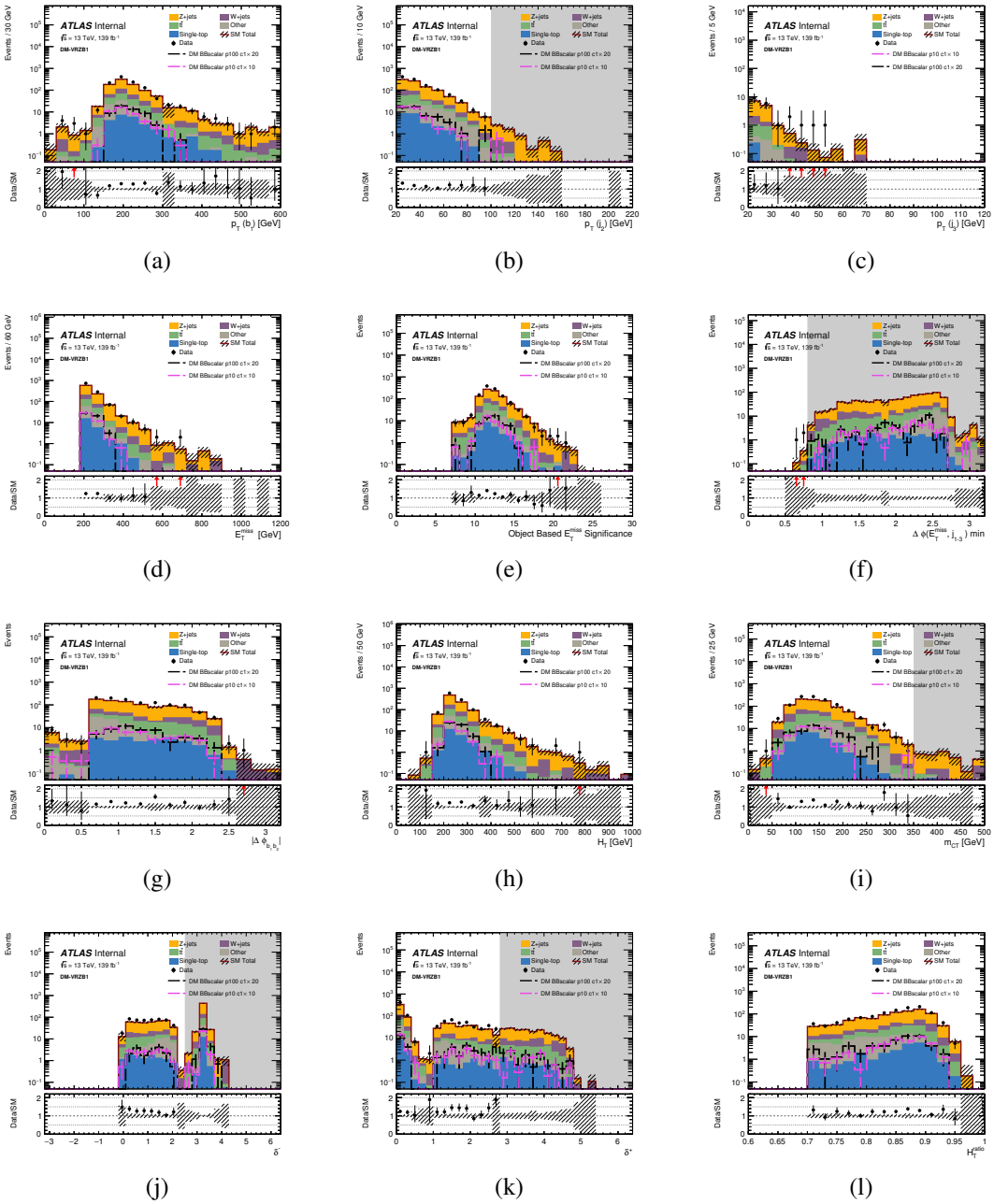


Fig. A.3 Data MC comparisons for all variables used in BDT training for DM-VRZB1.

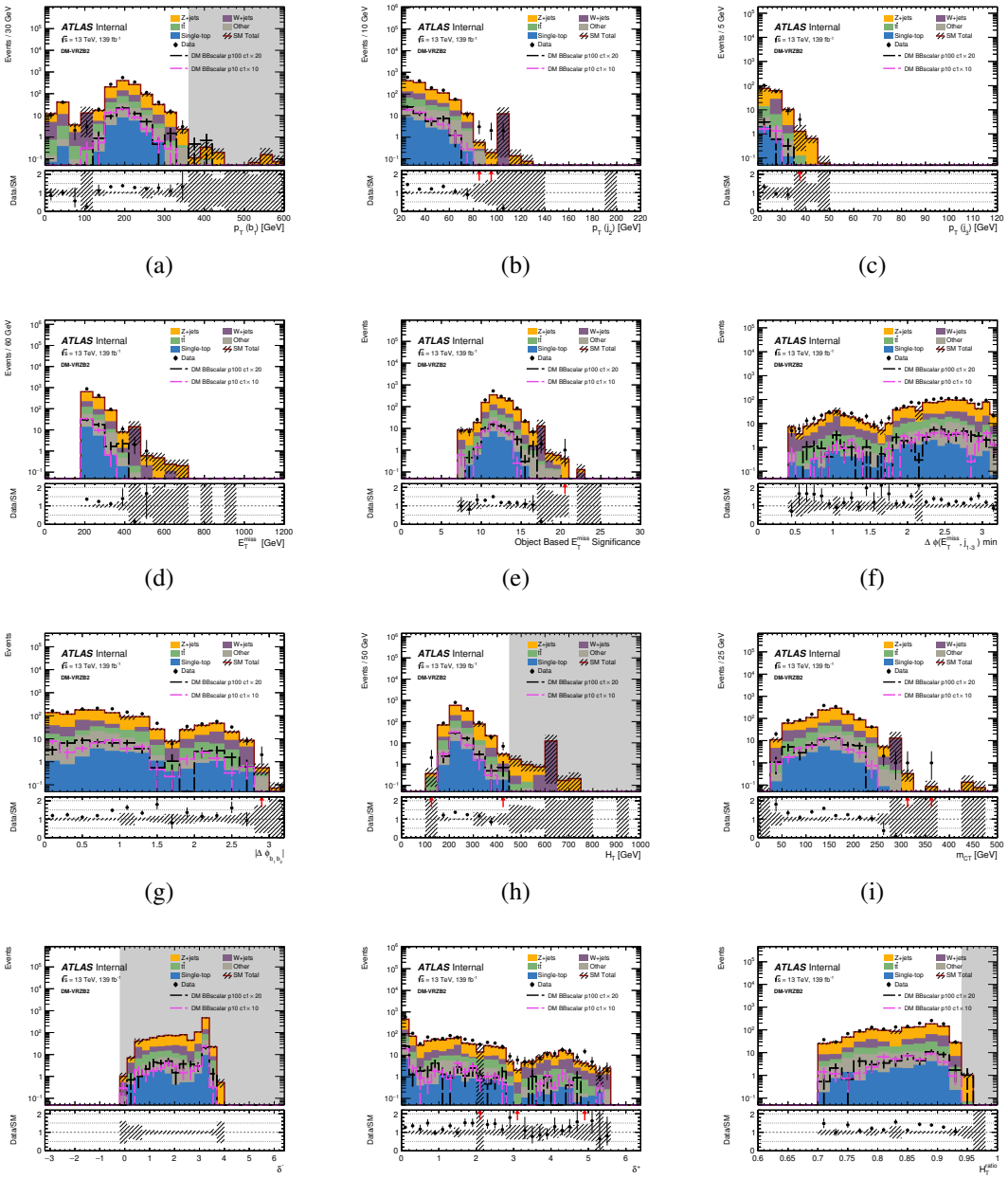
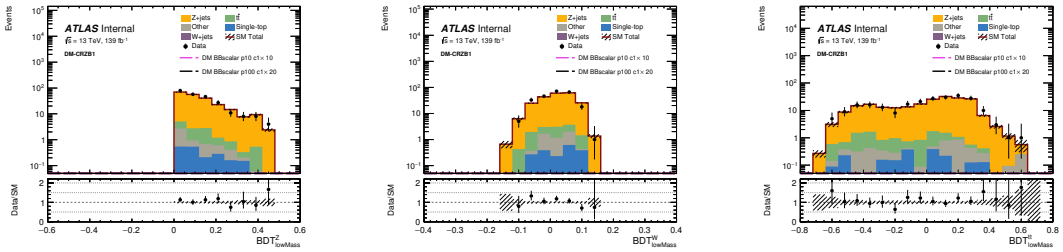
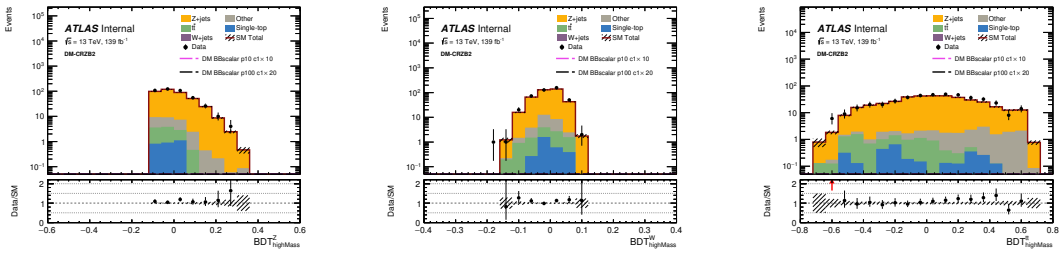


Fig. A.4 Data MC comparisons for all variables used in BDT training for DM-VRZB2.



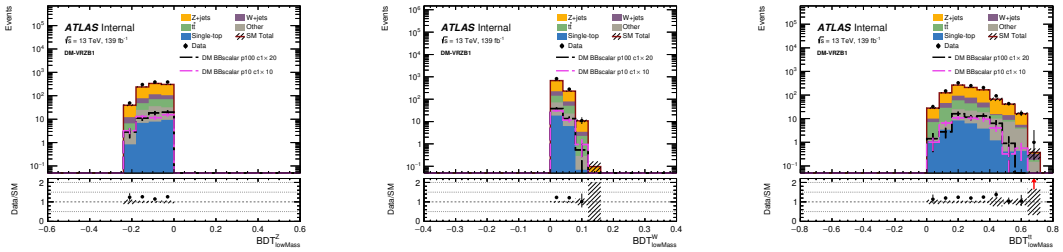
(a) (b) (c)

Fig. A.5 Data MC comparisons for all variables used in BDT training for DM-CRZB1.



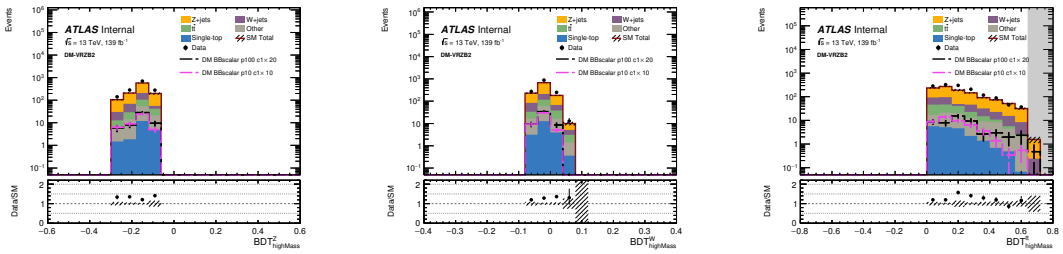
(a) (b) (c)

Fig. A.6 Data MC comparisons for all variables used in BDT training for DM-CRZB2.



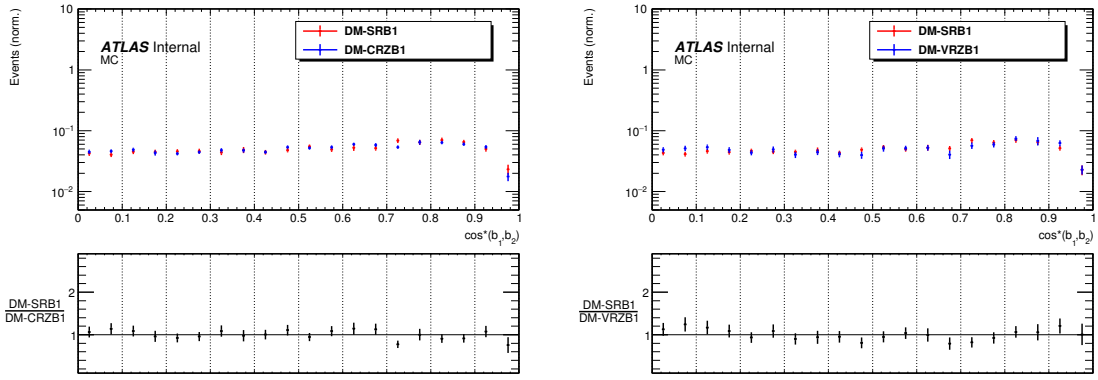
(a) (b) (c)

Fig. A.7 Data MC comparisons for all variables used in BDT training for DM-VRZB1.



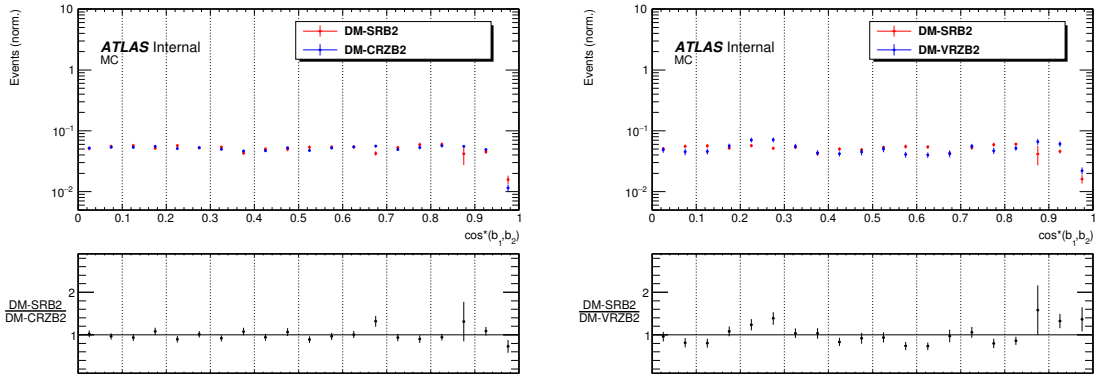
(a) (b) (c)

Fig. A.8 Data MC comparisons for all variables used in BDT training for DM-VRZB2.



(a)

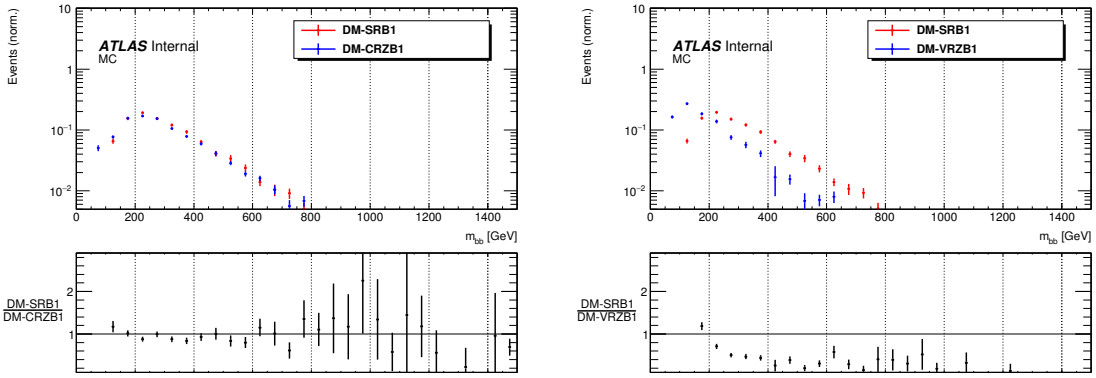
(b)



(c)

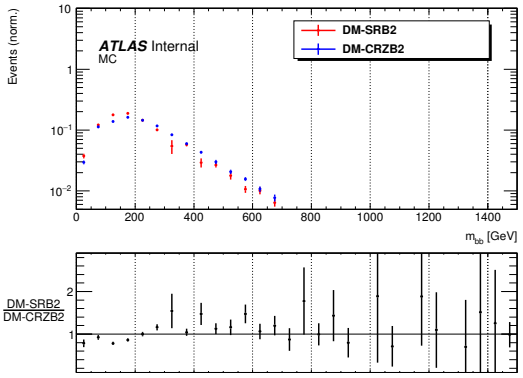
(d)

Fig. A.9 Shape comparisons between the CRs and VRs and the relevant SRs. To ensure that the difference in cuts does not affect the distribution of $\cos^*(b_1, b_2)$.

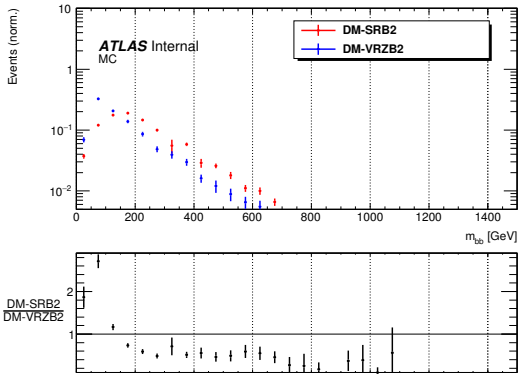


(a)

(b)

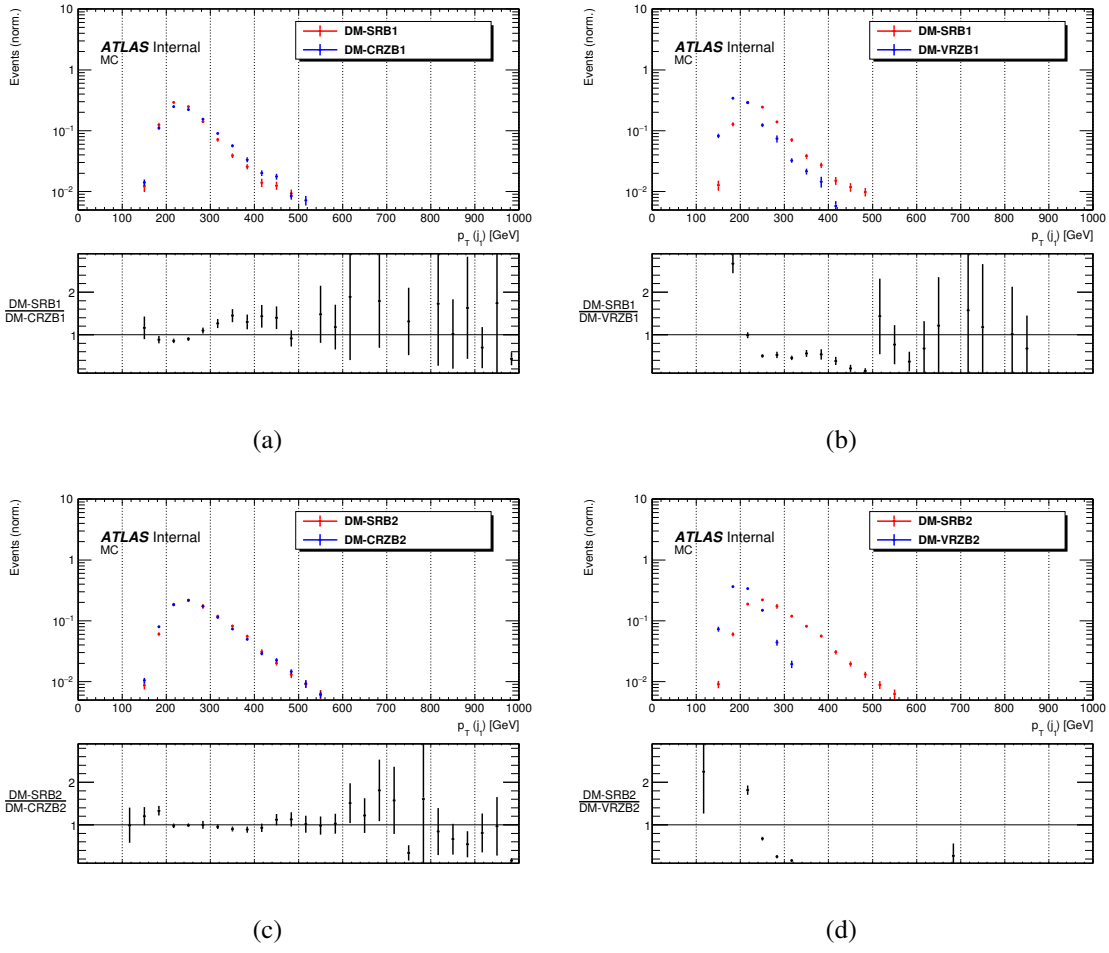


(c)



(d)

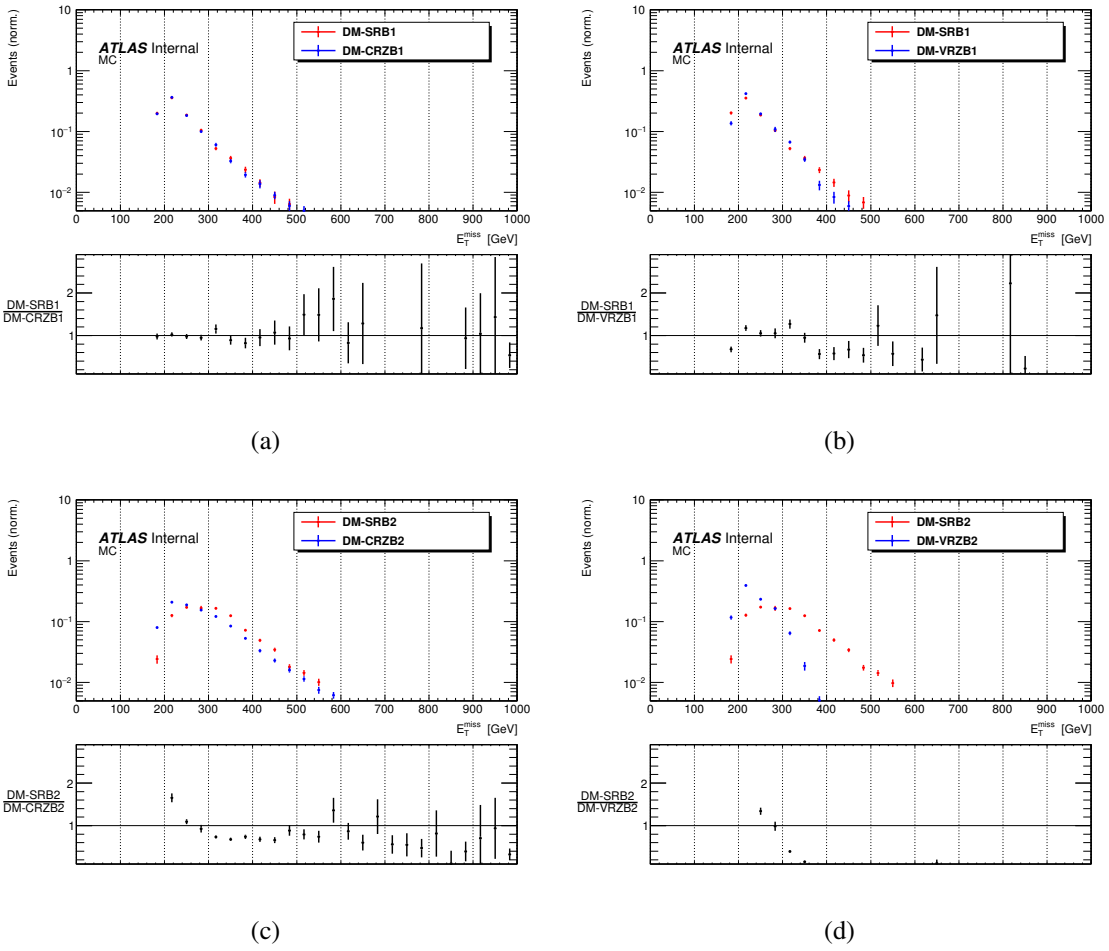
Fig. A.10 Shape comparisons between the CRs and VRs and the relevant SRs for m_{bb}.



(a) (b)

(c) (d)

Fig. A.11 Shape comparisons between the CRs and VRs and the relevant SRs for the p_T of the leading jet.



(a) (b)

(c) (d)

Fig. A.12 Shape comparisons between the CRs and VRs and the relevant SRs for E_T^{miss}.

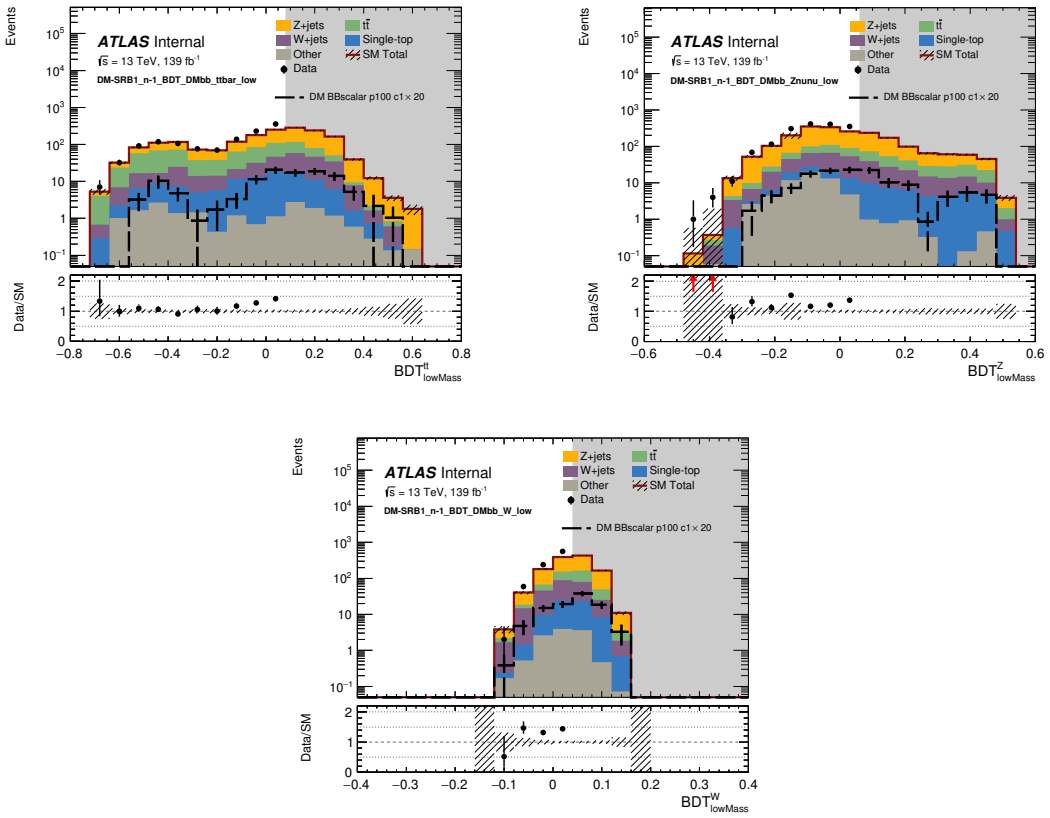


Fig. A.13 n-1 BDT distributions in DM-SRB1 for BDT discriminants trained on low mass DM mediator models.

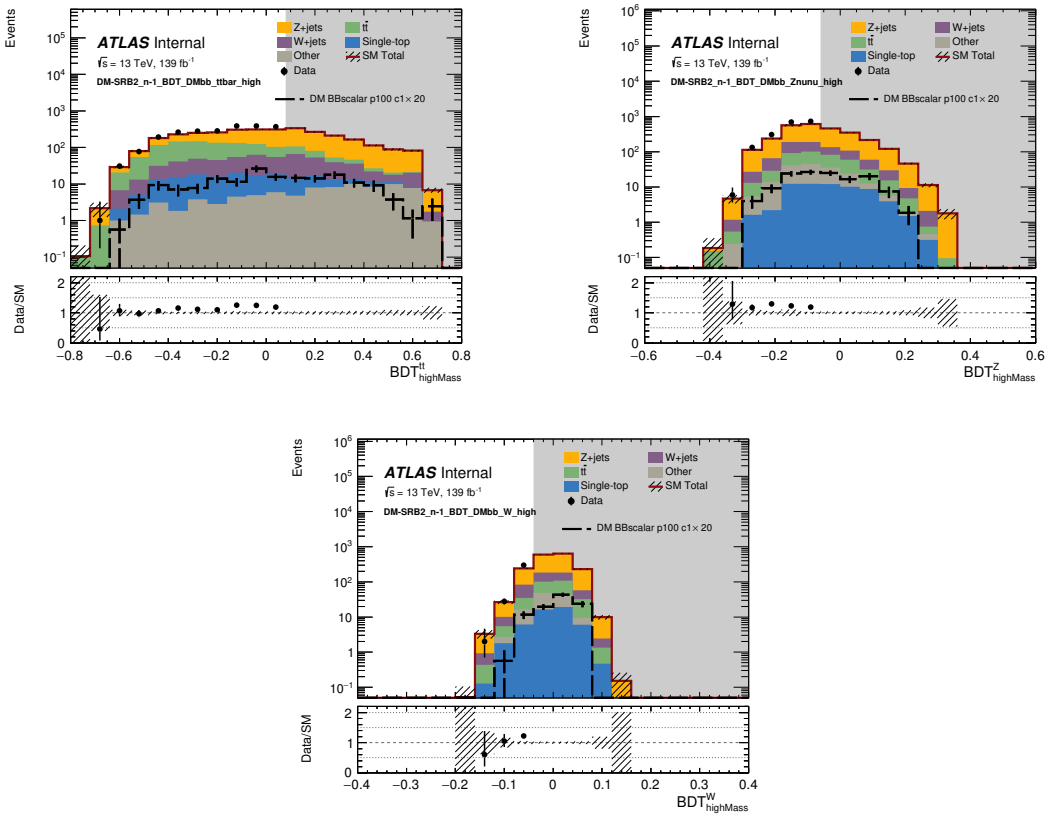


Fig. A.14 n-1 BDT distributions in DM-SRB2 for BDT discriminants trained on high mass DM mediator models.

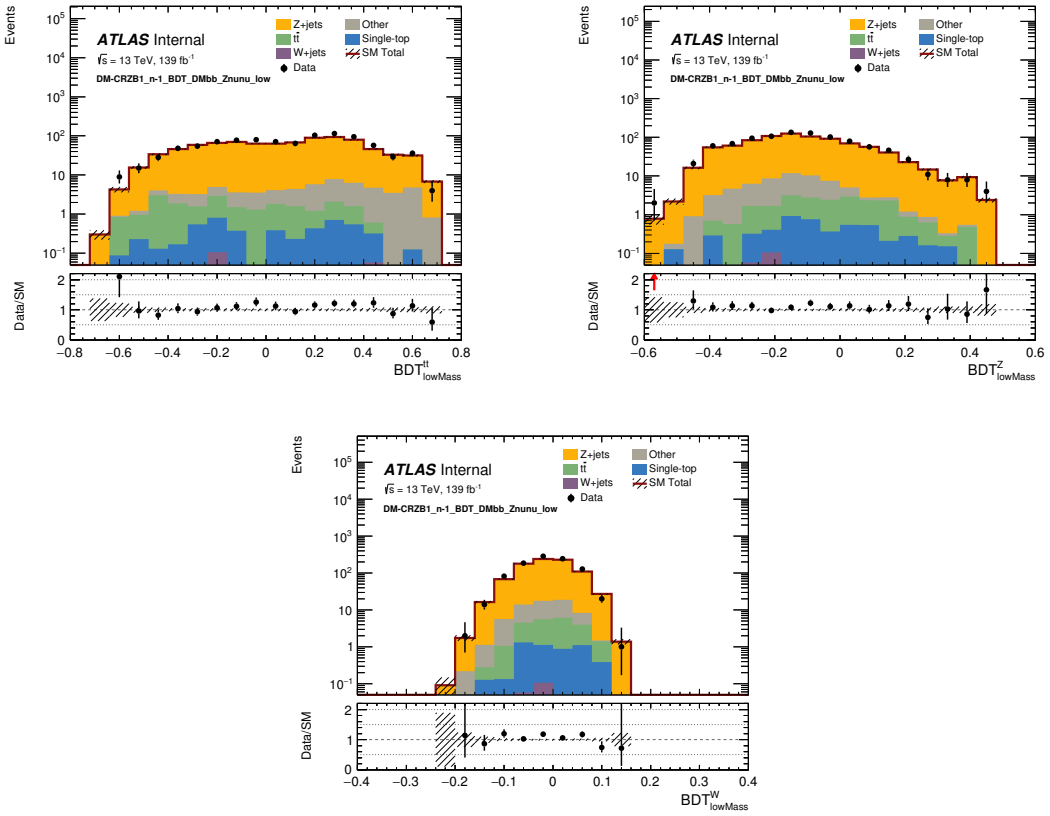


Fig. A.15 n-1 BDT distributions in DM-CRZB1 for BDT discriminants trained on low mass DM mediator models.

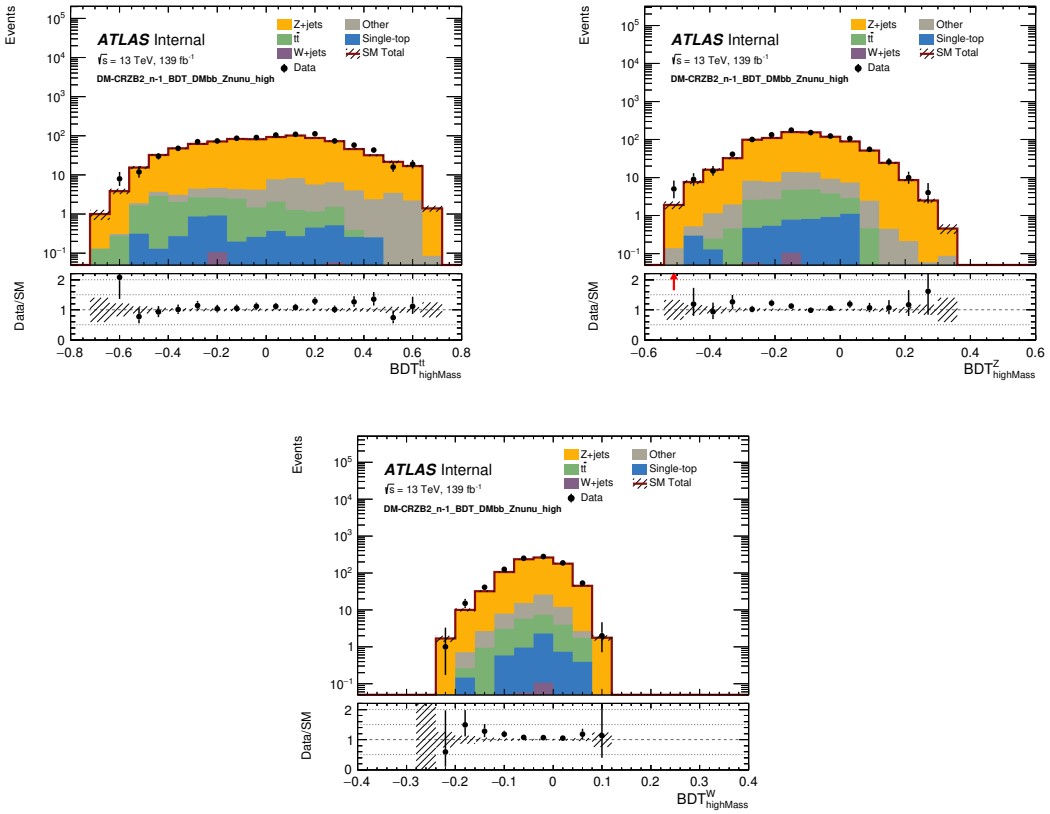


Fig. A.16 n-1 BDT distributions in DM-CRZB2 for BDT discriminants trained on high mass DM mediator models.

A.4 Signal Shape Comparison Plots

The following plots illustrate how the shape of the scalar and pseudo-scalar signal models vary for each of the candidate variables considered for the BDTs. One thing to note here is that there seems to be very little difference between scalar and pseudo-scalar signal models. The mass of the DM-SM mediator seems to have some limited influence on the shape of some distributions. For example some discrimination between high mass and low mass signal models can be seen in Figure A.20g and Figure A.20g as well as Figure A.21d and Figure A.21d. This is an indication that the scalar and pseudo-scalar MC datasets can be combined for BDT training, but it might be beneficial to perform separate trainings for high and low mass SM-DM mediator models.

A.5 Signal-Background Shape Comparison Plots

The following plots illustrate how the shape of signal the dominant analysis backgrounds vary compared to the combined high or low mass signal samples. These plots show that in general it is difficult to separate the $Z+$jets background from the signal models. Figure A.23a, Figure A.23b, Figure A.23g, Figure A.23h and Figure A.24c all show particularly prominent shape difference between the signal models and the backgrounds. This is an indication that these variables can be very powerful when trying to isolate the DM signals.

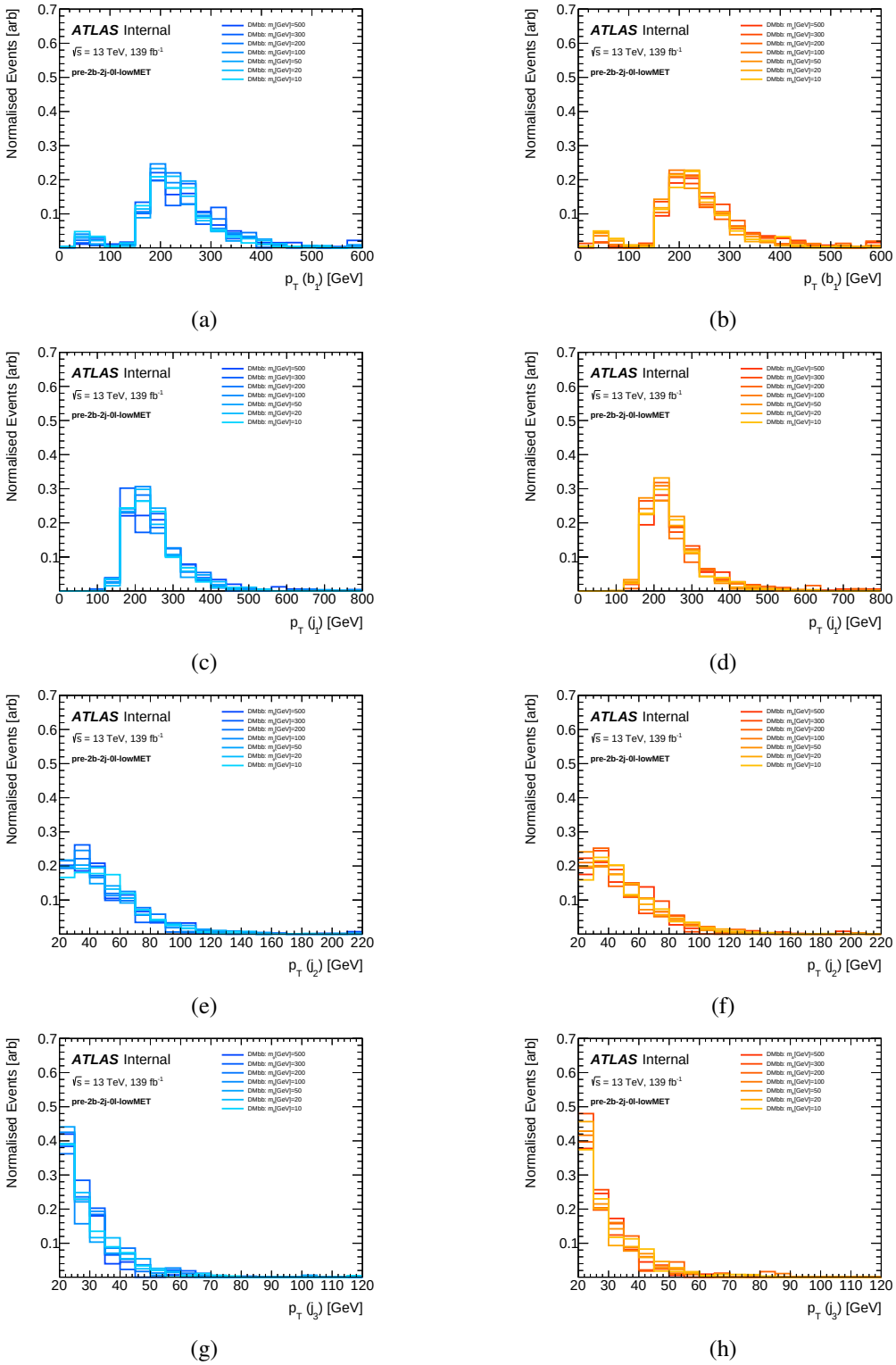


Fig. A.17 Signal comparison plots for $p_T(bj_1)$ ((a) and (b)), $p_T(j_1)$ ((c) and (d)), $p_T(j_2)$ ((e) and (f)) and $p_T(j_3)$ ((g) and (h)). The left hand plots are for scalar mediators and the right hand plots show pseudo-scalar mediators)

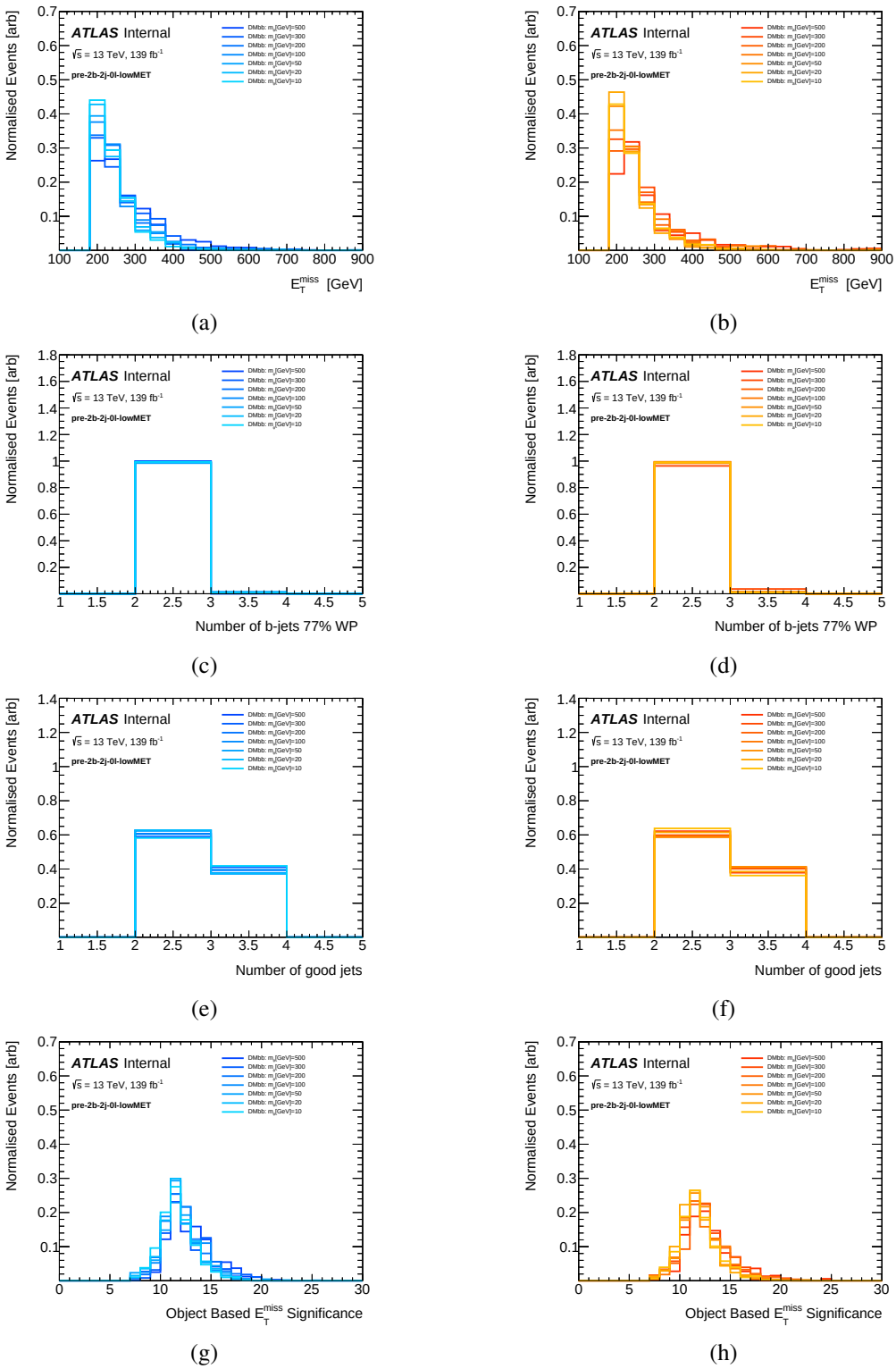


Fig. A.18 Signal comparison plots for E_T^{miss} ((a) and (b)), N_b ((c) and (d)), N_j^{good} ((e) and (f)) and S ((g) and (h)). The left hand plots are for scalar mediators and the right hand plots show pseudo-scalar mediators)

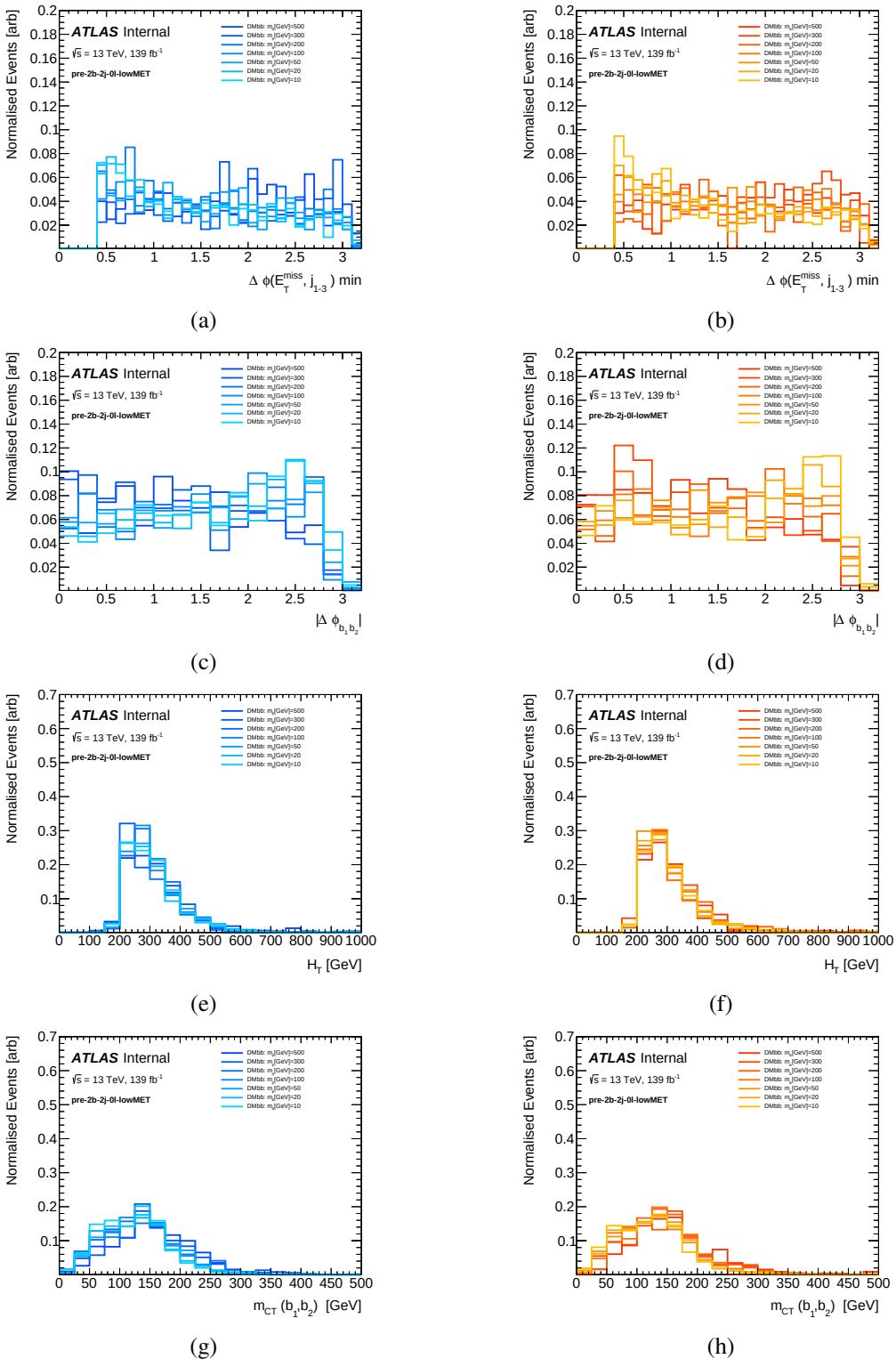


Fig. A.19 Signal comparison plots for $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$ ((a) and (b)), $\Delta\phi_{bb}$ ((c) and (d)), HT ((e) and (f)) and $\text{m}_{\text{CT}}(b_1, b_2)$ ((g) and (h)). The left hand plots are for scalar mediators and the right hand plots show pseudo-scalar mediators

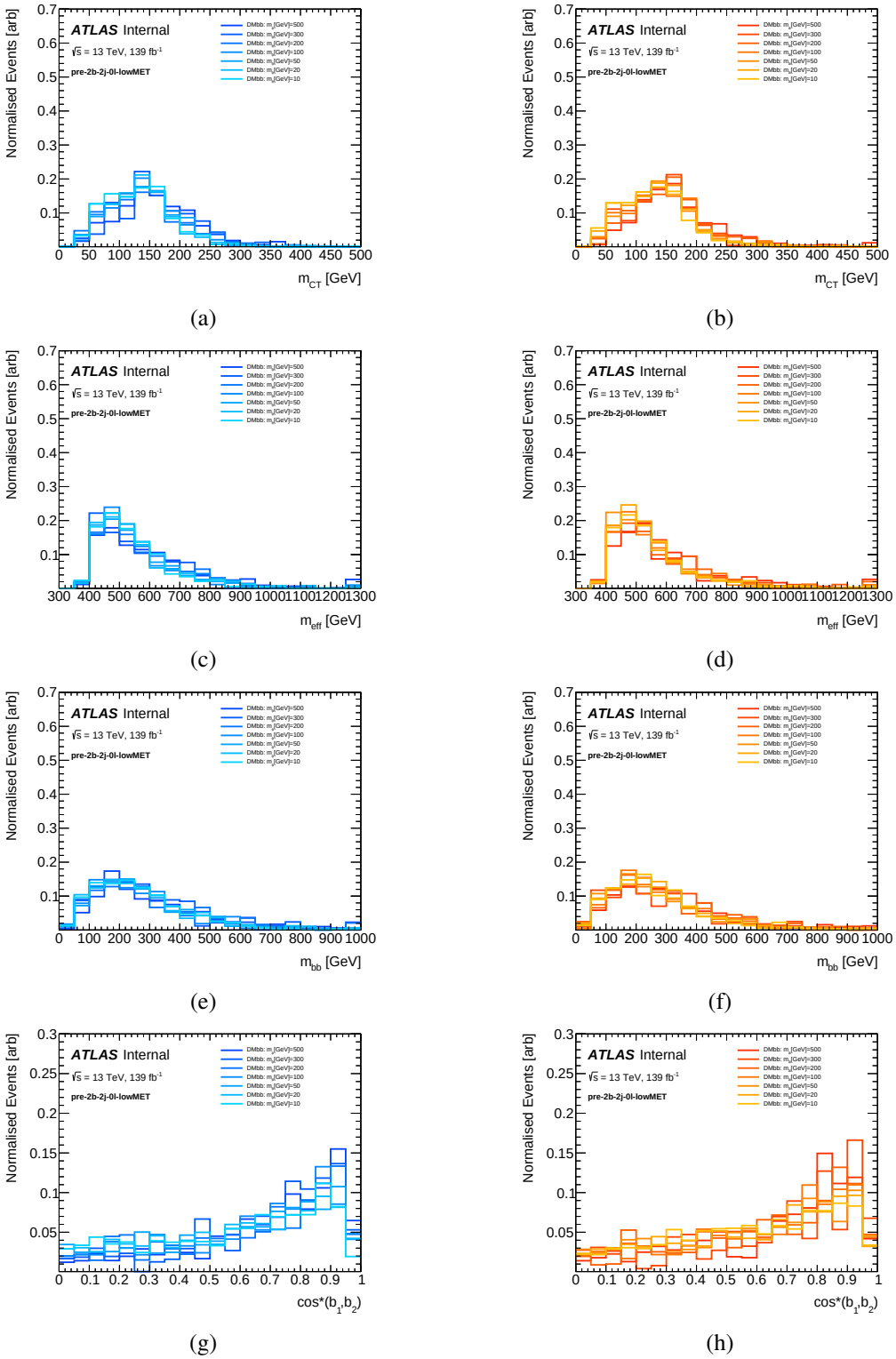


Fig. A.20 Signal comparison plots for $m_{CT}(j_1, j_2)$ ((a) and (b)), m_{eff} ((c) and (d)), m_{bb} ((e) and (f)) and $\cos^*(b_1, b_2)$ ((g) and (h)). The left hand plots are for scalar mediators and the right hand plots show pseudo-scalar mediators

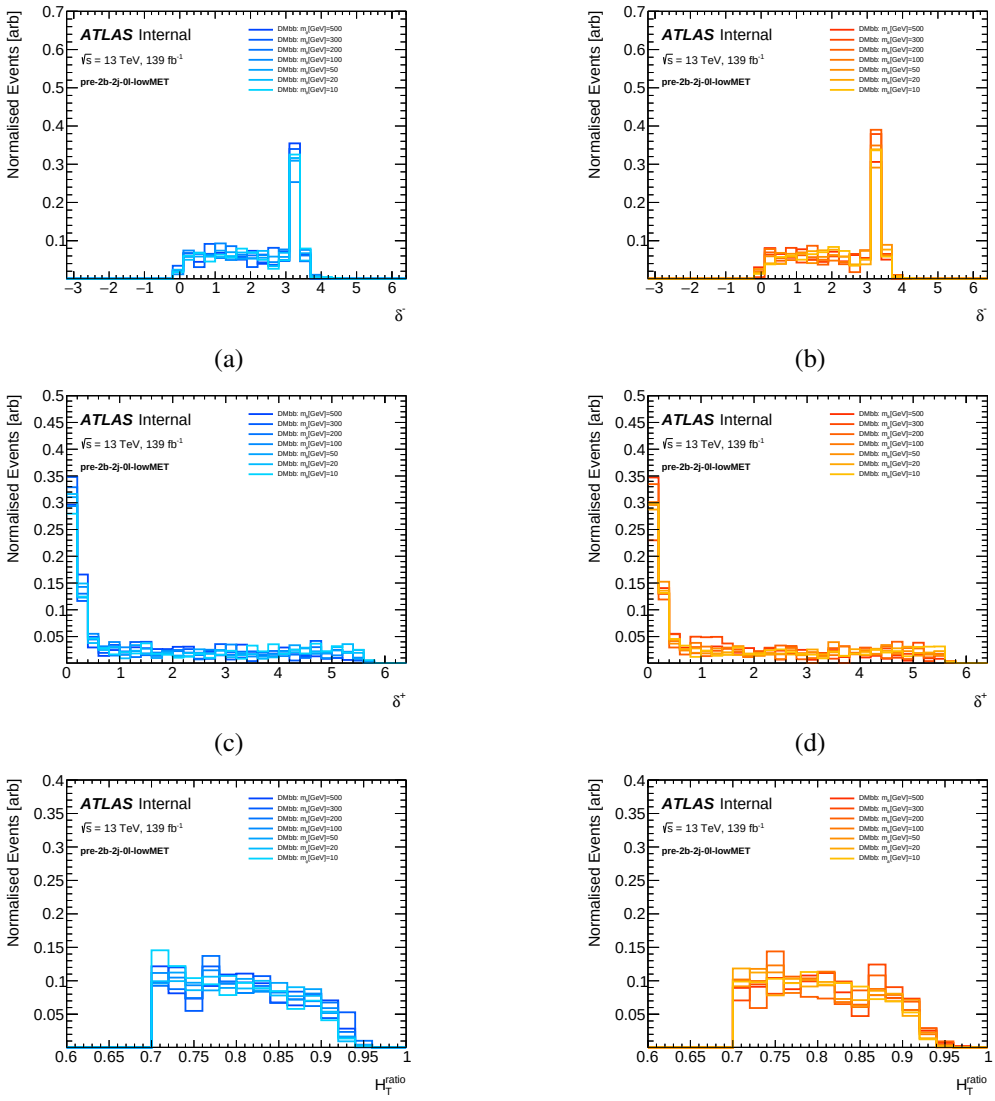
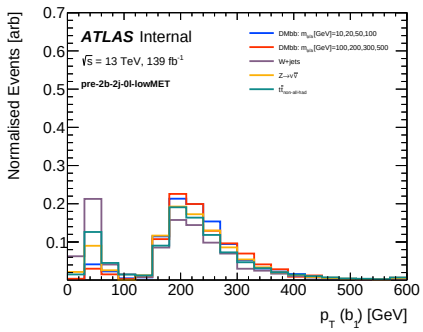
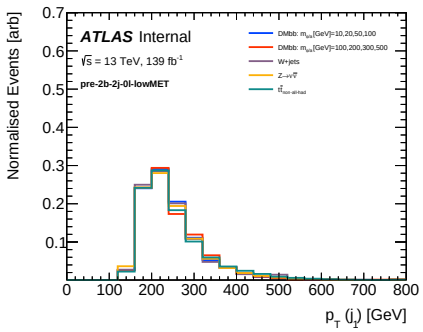


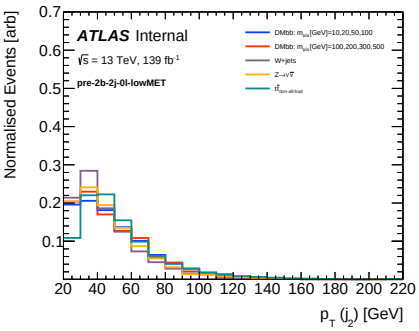
Fig. A.21 Signal comparison plots for δ^- ((a) and (b)), δ^+ ((c) and (d)) and H^T_{ratio} ((e) and (f)). The left hand plots are for scalar mediators and the right hand plots show pseudo-scalar mediators

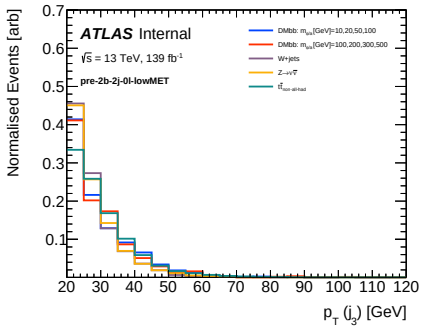


(a)

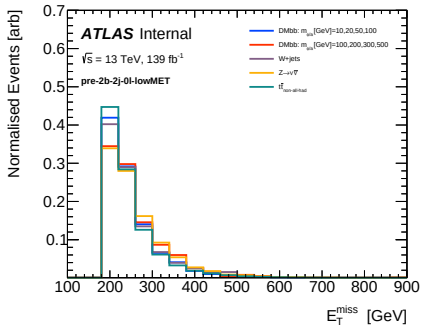


(b)

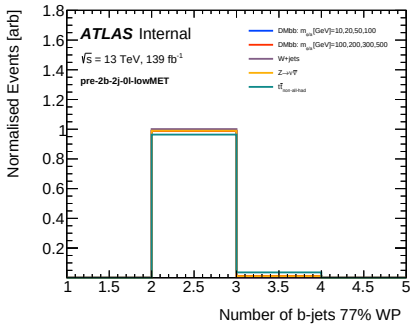




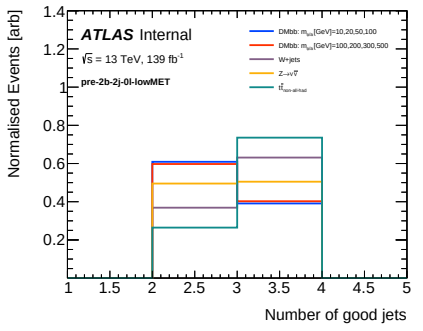
(c)



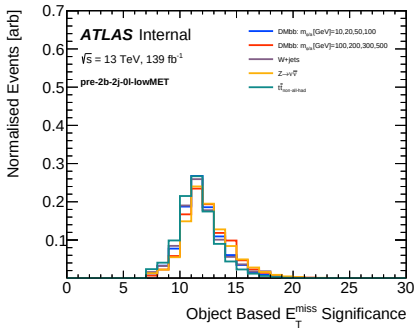
(d)



(e)



(f)



(g)

Fig. A.22 Signal vs background comparison plots for $p_T(bj_1)$ (a), $p_T(j_1)$ (b), $p_T(j_2)$ (c), $p_T(j_3)$ (d), E_T^{miss} (e), N_b (f), N_j^{good} (g) and S (h).

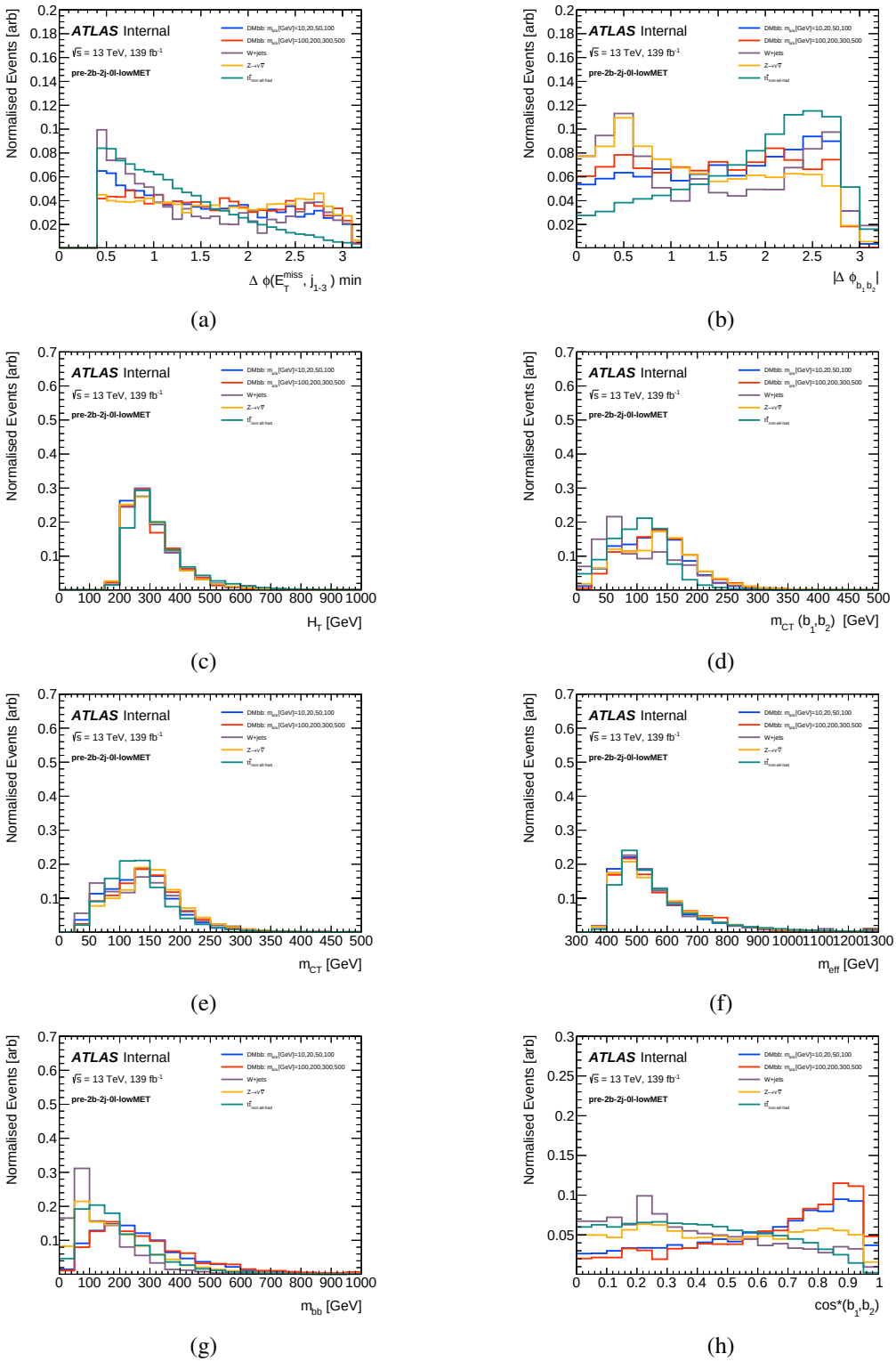
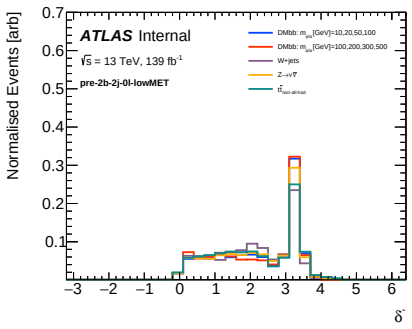
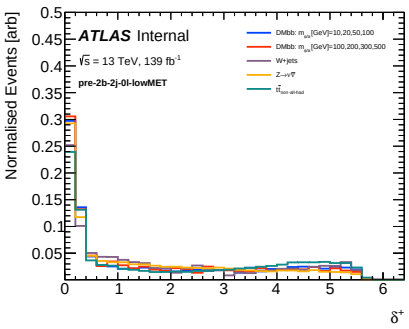


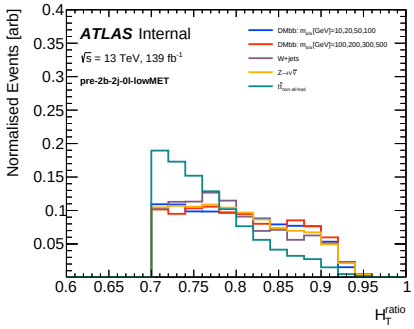
Fig. A.23 Signal vs background comparison plots for $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$ (a), $\Delta\phi_{bb}$ (b), HT (c) and $\text{m}_{\text{CT}}(b_1, b_2)$ (d) $\text{m}_{\text{CT}}(j_1, j_2)$ (e), m_{eff} (f), m_{bb} (g) and $\cos^*(b_1, b_2)$ (h).



(a)



(b)



(c)

Fig. A.24 Signal vs background comparison plots for δ^- (a), δ^+ (b) and H^T_{ratio} (c).

Appendix B

DM-SRA

This appendix will detail the definition and optimisation of a cut-based SR, DM-SRA, which is used to validate the BDT bases SRs.

DM-SRA has been designed to achieve optimal sensitivity for $b\bar{b}+\phi/a$ signals for all mediator masses ($m_{\phi/a}$ =10, 20, 50, 100, 200, 300, 500 GeV). These signals are characterised by two b-jets and comparatively low E_T^{miss}. The $\cos^*(b_1,b_2)$ variable has been found to be a powerful discriminating variable due to its sensitivity to the spin of the mediator [90]. Therefore the exclusion strategy of this analysis is to use a multi-bin fit in bins of $\cos^*(b_1,b_2)$.

E_T^{miss} Optimisation of DM-SRA In this section the optimisation of some basic cuts is performed for DM-SRA. The basis for these cuts is shown in Table B.1 and are taken from a paper studying the same signal models but only using 36 fb^{-1} of data [95]. Firstly a study is presented here as to whether the E_T^{miss} cut could be improved simultaneously with the $p_T(bj_1)$ and $p_T(j_1)$ cuts. For this study the $p_T(bj_1)$ and $p_T(j_1)$ cuts are kept equal to each other as was the case for the 36 fb^{-1} analysis. These are then iterated from 40 GeV to 240 GeV in intervals of 40 GeV. The E_T^{miss} cuts are iterated from 180 GeV to 300 GeV in intervals of 20 GeV. trigPlateau > 1 is always applied such that the signal region will be trigger safe even if the cut on E_T^{miss} is low.

Figures B.1, B.2 and B.3 show how an approximate limit on the cross section of the mediator changes accross the entire range of masses considered. This limit is naively calculated by multiplying the yield of the signal by a scale factor until 1σ signal significance is achieved (the lower edge of the shaded bands), then 2σ (the solid line in the middle of the shaded band) and 3σ (the upper edge of the shaded band). The significance is calculated using a binomial figure of merit and a flat systematic uncertainty of 30% is applied here. Figures B.1, B.2 and B.3 show signal significance for each mass point for both scalar and pseudo-scalar signal DM mediators while varying first the E_T^{miss} lower band between 180

Variable	Selection
Lowest unprescaled trigger	HLT_xe
SUSY1 Filtering	✓
Number of *baseline* leptons	0
E_T^{miss}[GeV]	> 180
Number of *signal* jets	$\in [2,3]$
Number of b-tagged jets	2
$p_{\text{T}}(j_1)$ [GeV]	> 150
$p_{\text{T}}(j_2)$ [GeV]	> 150
$p_{\text{T}}(j_3)$ [GeV]	< 60
$\Delta\phi(E_T^{\text{miss}}, j^{1-3})$	> 0.4
H_{ratio}^T	> 0.75
δ^-	< 0
δ^+	< 0.5

Table B.1 Definition of the DM-SRA from paper using 36 fb^{-1} of data [95].

and 300 GeV in steps of 20 GeV and then varying the $p_T(j_1)$ lower bound between 40 and 240 GeV in steps of 40 GeV. From the plots it is clear that varying $p_T(j_1)$ does not have a significant impact on the significance. There is a slight improvement in sensitivity associated with a E_T^{miss} cut of 180 GeV. As well as this, it can be seen from the instability of the significance calculations for higher E_T^{miss} cuts that there is a significant decrease in statistics in this phase space.

Also considered were varying cuts on the trigPlateau variable. A minimum value of trigPlateau>1 is required in order to ensure that the trigger is fully efficient in the region of phase space being considered. Figure B.4 shows the limit on the cross section (i.e. $\frac{\sigma}{sigma(g=1)}$) of the mediator for all signal masses varying the cut on trigPlateau between 1 and 10 in steps of 1. It is found that for low mediator masses, where the best sensitivity is achieved, the cut that yields the best sensitivity is trigPlateau>1.

B.2 shows the expected yields for E_T^{miss} cuts of 200 GeV and 240 GeV respectively. This illustrates how having a lower E_T^{miss} cut allows for a higher signal efficiency. The last column of this table also shows the expected yields for DM-SRA but using trigPlateau>1 and shows how this yields a higher signal efficiency still.

From this it can generally be seen that using trigPlateau>1 yields a comparably good sensitivity to any set of 1D cuts. The advantage of using the trigPlateau cut instead is that the signal efficiency is higher and therefore there are more statistics available.

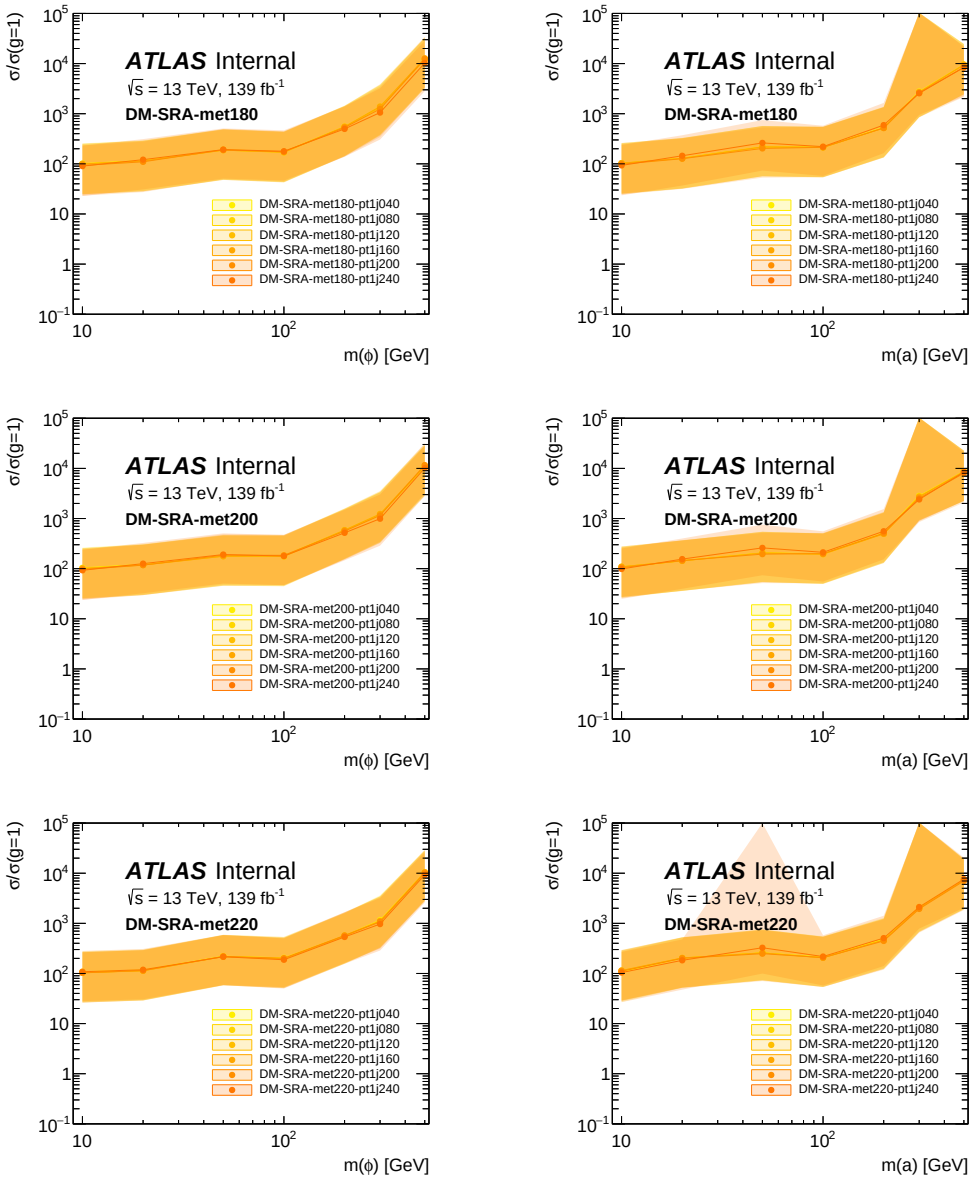


Fig. B.1 Sensitivity for different E_T^{miss},$p_T(bj_1)$ and $p_T(j_1)$ cuts.

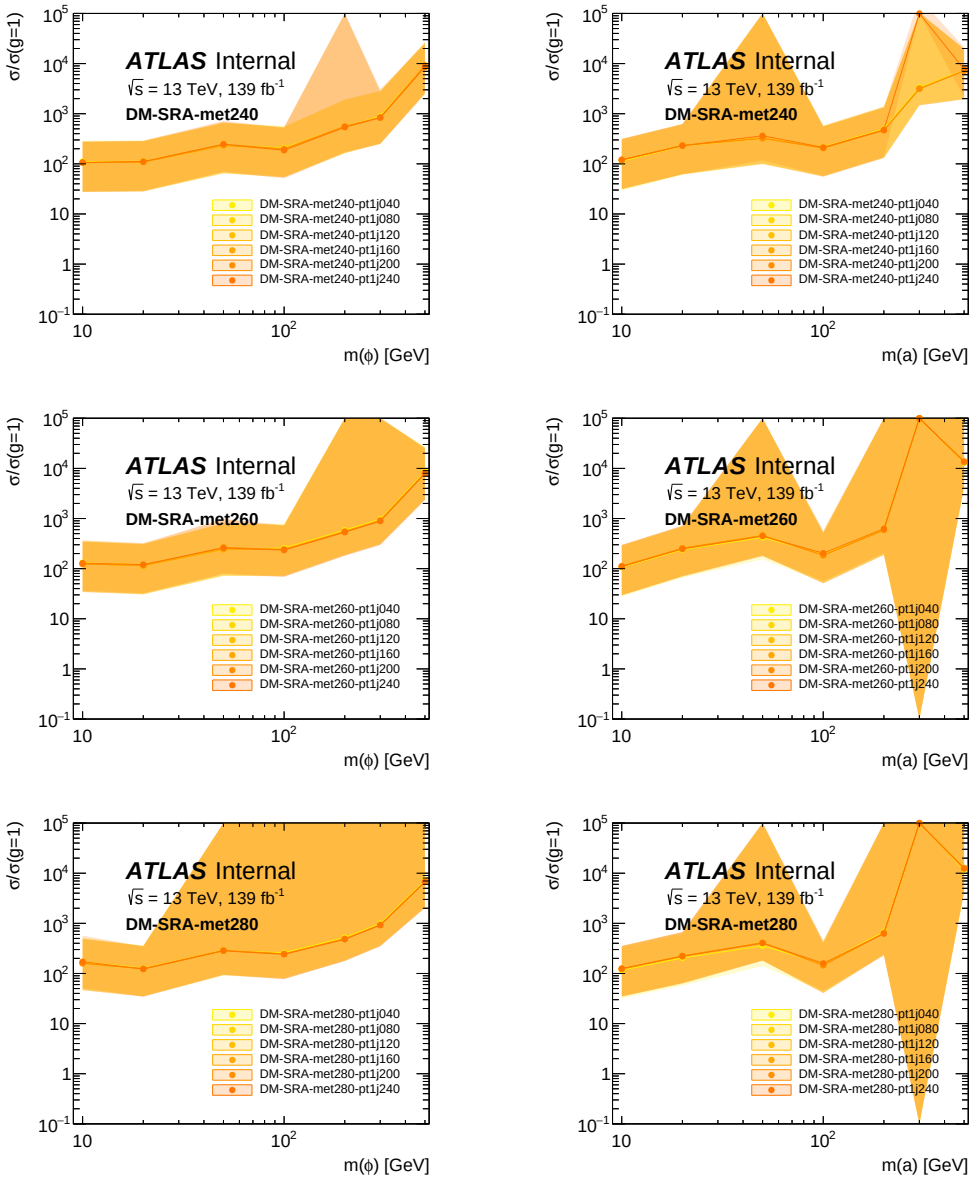


Fig. B.2 Sensitivity for different E_T^{miss},$p_T(bj_1)$ and $p_T(j_1)$ cuts.

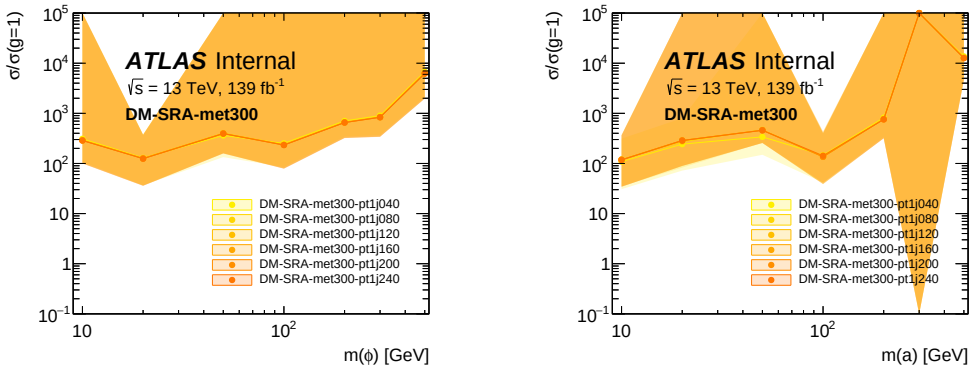


Fig. B.3 Sensitivity for different E_T^{miss},$p_T(bj_1)$ and $p_T(j_1)$ cuts.

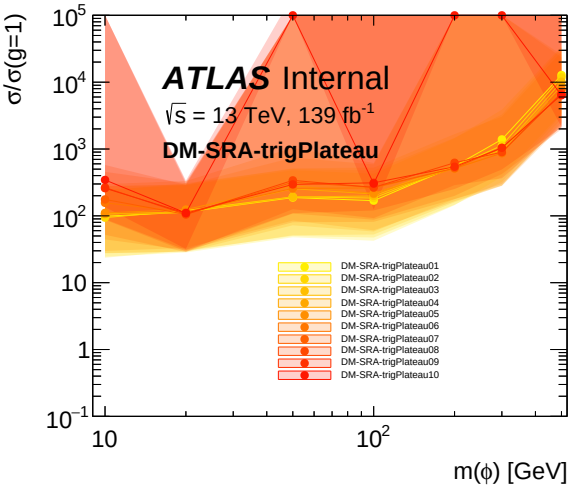


Fig. B.4 Sensitivity for different trigPlateau cuts.

	$E_T^{\text{miss}} > 200$ GeV $p_{\text{T}}(j_1) > 160$ GeV	$E_T^{\text{miss}} > 240$ GeV $p_{\text{T}}(j_1) > 160$ GeV	trigPlateau > 1
MC exp. Z events	564.80 ± 18.33	307.49 ± 12.75	720.25 ± 21.63
MC exp. ttbar events	219.37 ± 4.45	111.58 ± 2.17	278.94 ± 5.77
MC exp. W events	166.35 ± 16.62	86.11 ± 12.60	210.72 ± 19.70
MC exp. st events	39.26 ± 1.99	20.68 ± 1.53	49.92 ± 2.24
MC exp. other events	9.47 ± 1.45	5.71 ± 1.19	5.64 ± 5.63
$m(\phi,\chi) = (10,1)$ GeV $, g = 1$	4.87 ± 0.53	2.56 ± 0.39	6.21 ± 0.58
$m(a,\chi) = (10,1)$ GeV $, g = 1$	4.47 ± 0.48	2.23 ± 0.35	5.92 ± 0.54
$m(\phi,\chi) = (20,1)$ GeV $, g = 1$	4.13 ± 0.45	2.47 ± 0.37	5.38 ± 0.51
$m(a,\chi) = (20,1)$ GeV $, g = 1$	3.41 ± 0.37	1.17 ± 0.21	4.75 ± 0.44
$m(\phi,\chi) = (50,1)$ GeV $, g = 1$	2.81 ± 0.40	1.22 ± 0.26	3.32 ± 0.44
$m(a,\chi) = (50,1)$ GeV $, g = 1$	2.47 ± 0.42	0.94 ± 0.26	2.91 ± 0.45
$m(\phi,\chi) = (100,1)$ GeV $, g = 1$	2.74 ± 0.35	1.41 ± 0.28	3.65 ± 0.40
$m(a,\chi) = (100,1)$ GeV $, g = 1$	2.51 ± 0.33	1.31 ± 0.24	2.88 ± 0.35
$m(\phi,\chi) = (200,1)$ GeV $, g = 1$	0.87 ± 0.16	0.54 ± 0.14	1.15 ± 0.18
$m(a,\chi) = (200,1)$ GeV $, g = 1$	1.01 ± 0.16	0.60 ± 0.13	1.21 ± 0.17
$m(\phi,\chi) = (300,1)$ GeV $, g = 1$	0.43 ± 0.10	0.35 ± 0.09	0.47 ± 0.10
$m(a,\chi) = (300,1)$ GeV $, g = 1$	0.23 ± 0.07	0.14 ± 0.05	0.27 ± 0.08
$m(\phi,\chi) = (500,1)$ GeV $, g = 1$	0.05 ± 0.01	0.03 ± 0.01	0.05 ± 0.01
$m(a,\chi) = (500,1)$ GeV $, g = 1$	0.06 ± 0.01	0.04 ± 0.01	0.07 ± 0.01
MC exp. SM events	999.26 ± 25.26	531.57 ± 18.16	1265.47 ± 30.42

Table B.2 Expected MC Yields for a preliminary definition of DM-SRA with either $E_T^{\text{miss}} > 200$ GeV and $p_{\text{T}}(j_1) > 160$ GeV, $E_T^{\text{miss}} > 240$ GeV and $p_{\text{T}}(j_1) > 160$ GeV or trigPlateau > 1.

δ^- and H^T_{ratio} Optimisation for DM-SRA Also studied are the optimisations on the H^T_{ratio} and δ^- variables. Tables B.3, B.4, B.5 and B.6 show the yields for DM-SRA first applying a low limit on H^T_{ratio} between 0.7 and 0.85 in steps of 0.5, then applying upper limit cuts on δ^- between 1.5 and -1.5 in steps of 0.5. Cuts of $H^T_{\text{ratio}} > 0.8$ and $\delta^- < 0$ were chosen in order to optimise sensitivity whilst maintaining signal efficiency.

Cut Applied	TTBar	Single Top	DiJet	Z+Jets	W+Jets	DMbb p20 *200	significance
passMETtriggers	325779.9	39437.8	601773.5	204080.1	321431.8	324803.0	0.488
$\Delta\Phi(E_T^{\text{miss}}, j^{1-4}) > 0.4$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} >= 2$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} <= 3$	58957.5	9893.8	10995.2	108916.1	122902.2	187896.0	1.527
$N_b >= 2$	16285.1	2202.3	71.7	9925.2	6908.1	37834.7	2.586
$N_l^{\text{baseline}} == 0$	12182.5	1196.2	70.8	9319.3	5712.5	37834.7	3.111
$E_T^{\text{miss}} > \frac{5000}{p_T(j_1)-20} + 160$	11596.0	1150.7	70.8	8925.2	5471.4	36580.8	3.142
$p_T(bj_1) > 150$	6290.7	730.6	70.8	5596.5	3074.7	24022.9	3.483
$p_T(j_3) < 60$	4473.2	597.1	25.0	5258.9	2786.5	21592.5	3.7
$H^T_{\text{ratio}} > 0.7$	1786.4	316.9	1.9	3426.3	1711.6	13696.9	4.134
$\delta^- < 1.5$	835.7	138.1	1.9	1225.0	659.2	5837.2	4.381
$\delta^- < 1$	766.7	125.2	1.9	1030.5	559.1	5113.2	4.41
$\delta^- < 0.5$	688.5	112.4	1.9	839.4	477.1	4344.0	4.393
$\delta^- < 0$	600.5	98.2	1.9	668.4	378.7	3739.4	4.539
$\delta^- < -0.5$	494.6	81.9	1.9	516.0	312.9	3128.1	4.669
$\delta^- < -1$	372.2	66.1	1.9	366.9	246.9	2312.2	4.618
$\delta^- < -1.5$	242.7	43.4	1.9	231.2	172.0	1701.8	5.027

Table B.3 Cut flow for DM-SRA with a cut of $H^T_{\text{ratio}} > 0.7$ and varying upper limit cuts on δ^- between 1.5 and -1.5 in steps of 0.5.

Cut Applied	TTBar	Single Top	DiJet	Z+Jets	W+Jets	DMbb p20 *200	significance
passMETtriggers	325779.9	39437.8	601773.5	204080.1	321431.8	324803.0	0.488
$\Delta\Phi(E_T^{\text{miss}}, j^{1-4}) > 0.4$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} >= 2$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} <= 3$	58957.5	9893.8	10995.2	108916.1	122902.2	187896.0	1.527
$N_b >= 2$	16285.1	2202.3	71.7	9925.2	6908.1	37834.7	2.586
$N_l^{\text{baseline}} == 0$	12182.5	1196.2	70.8	9319.3	5712.5	37834.7	3.111
$E_T^{\text{miss}} > \frac{5000}{p_T(j_1)-20} + 160$	11596.0	1150.7	70.8	8925.2	5471.4	36580.8	3.142
$p_T(bj_1) > 150$	6290.7	730.6	70.8	5596.5	3074.7	24022.9	3.483
$p_T(j_3) < 60$	4473.2	597.1	25.0	5258.9	2786.5	21592.5	3.7
$H^T_{\text{ratio}} > 0.75$	1024.2	209.5	1.9	2612.1	1288.9	10392.9	4.355
$\delta^- < 1.5$	476.0	87.1	1.9	894.4	472.8	4326.1	4.699
$\delta^- < 1$	436.0	79.5	1.9	764.8	401.8	3802.3	4.727
$\delta^- < 0.5$	390.0	70.3	1.9	630.4	340.3	3205.8	4.693
$\delta^- < 0$	340.6	61.2	1.9	505.4	266.5	2772.0	4.879
$\delta^- < -0.5$	282.2	50.0	1.9	394.1	222.8	2255.5	4.897
$\delta^- < -1$	208.2	39.2	1.9	282.7	177.9	1676.1	4.873
$\delta^- < -1.5$	132.6	25.2	1.9	178.4	125.8	1259.9	5.391

Table B.4 Cut flow for DM-SRA with a cut of $H^T_{\text{ratio}} > 0.75$ and varying upper limit cuts on δ^- between 1.5 and -1.5 in steps of 0.5.

Cut Applied	TTBar	Single Top	DiJet	Z+Jets	W+Jets	DMbb p20 *200	significance
passMETtriggers	325779.9	39437.8	601773.5	204080.1	321431.8	324803.0	0.488
$\Delta\Phi(E_T^{\text{miss}}, j^{1-4}) > 0.4$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} >= 2$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} <= 3$	58957.5	9893.8	10995.2	108916.1	122902.2	187896.0	1.527
$N_b >= 2$	16285.1	2202.3	71.7	9925.2	6908.1	37834.7	2.586
$N_l^{\text{baseline}} == 0$	12182.5	1196.2	70.8	9319.3	5712.5	37834.7	3.111
$E_T^{\text{miss}} > \frac{5000}{p_T(j_1)-20} + 160$	11596.0	1150.7	70.8	8925.2	5471.4	36580.8	3.142
$p_T(bj_1) > 150$	6290.7	730.6	70.8	5596.5	3074.7	24022.9	3.483
$p_T(j_3) < 60$	4473.2	597.1	25.0	5258.9	2786.5	21592.5	3.7
$H_{\text{ratio}}^T > 0.8$	498.6	111.9	0.0	1727.7	857.5	6870.4	4.56
$\delta^- < 1.5$	219.5	42.3	0.0	551.2	300.0	2749.0	5.05
$\delta^- < 1$	199.3	37.9	0.0	484.9	262.5	2439.5	5.059
$\delta^- < 0.5$	177.8	33.4	0.0	405.9	229.0	2063.5	4.997
$\delta^- < 0$	153.5	28.7	0.0	330.8	184.0	1823.7	5.259
$\delta^- < -0.5$	123.7	24.1	0.0	261.3	157.5	1524.0	5.359
$\delta^- < -1$	89.6	18.8	0.0	190.1	124.6	1126.2	5.306
$\delta^- < -1.5$	56.5	11.8	0.0	122.6	93.6	760.8	5.299

Table B.5 Cut flow for DM-SRA with a cut of $H_{\text{ratio}}^T > 0.8$ and varying upper limit cuts on δ^- between 1.5 and -1.5 in steps of 0.5.

Cut Applied	TTBar	Single Top	DiJet	Z+Jets	W+Jets	DMbb p20 *200	significance
passMETtriggers	325779.9	39437.8	601773.5	204080.1	321431.8	324803.0	0.488
$\Delta\Phi(E_T^{\text{miss}}, j^{1-4}) > 0.4$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} >= 2$	237159.5	26351.3	27375.5	169766.0	221286.3	257639.8	0.943
$N_j^{\text{good}} <= 3$	58957.5	9893.8	10995.2	108916.1	122902.2	187896.0	1.527
$N_b >= 2$	16285.1	2202.3	71.7	9925.2	6908.1	37834.7	2.586
$N_l^{\text{baseline}} == 0$	12182.5	1196.2	70.8	9319.3	5712.5	37834.7	3.111
$E_T^{\text{miss}} > \frac{5000}{p_T(j_1)-20} + 160$	11596.0	1150.7	70.8	8925.2	5471.4	36580.8	3.142
$p_T(bj_1) > 150$	6290.7	730.6	70.8	5596.5	3074.7	24022.9	3.483
$p_T(j_3) < 60$	4473.2	597.1	25.0	5258.9	2786.5	21592.5	3.7
$H_{\text{ratio}}^T > 0.85$	204.8	52.4	0.0	981.2	478.5	3483.1	4.357
$\delta^- < 1.5$	86.4	19.6	0.0	295.8	155.6	1259.4	4.706
$\delta^- < 1$	76.8	16.9	0.0	275.0	134.6	1097.5	4.577
$\delta^- < 0.5$	67.8	14.7	0.0	229.4	117.3	933.8	4.563
$\delta^- < 0$	56.6	12.3	0.0	188.1	99.6	788.9	4.612
$\delta^- < -0.5$	44.9	10.2	0.0	147.8	84.0	615.9	4.494
$\delta^- < -1$	33.5	7.6	0.0	109.1	64.6	478.7	4.605
$\delta^- < -1.5$	21.9	4.7	0.0	67.7	47.6	291.2	4.281

Table B.6 Cut flow for DM-SRA with a cut of $H_{\text{ratio}}^T > 0.85$ and varying upper limit cuts on δ^- between 1.5 and -1.5 in steps of 0.5.

DM-SRA Definition Table B.7 shows the optimised cuts for DM-SRA

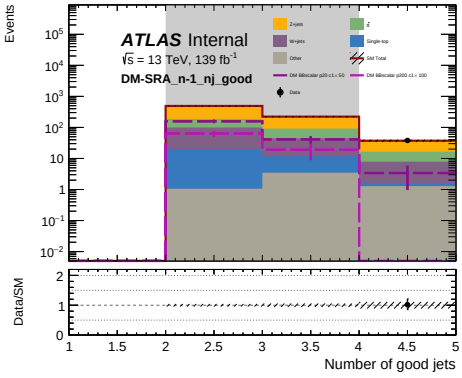
Variable	DM-SRA
DM-pre-2b-2j3-0l	
$p_T(bj_1)$	> 150 GeV
$p_T(j_3)$	< 60 GeV
$p_T(j_3)$	< 60 GeV
H^T_{ratio}	> 0.8
δ^-	< 0

Table B.7 Summary of the selections that define DM-SRA .

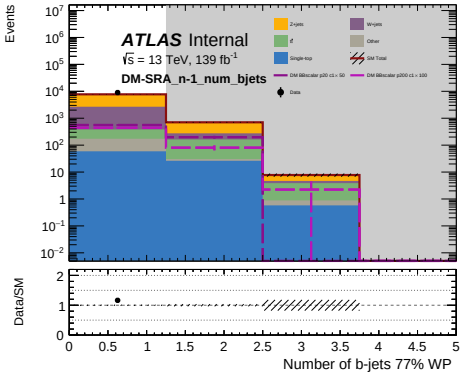
Table B.8 presents pre-fit expected yields for DM-SRA. Based on the yields here, a CR will be defined for W+jets, $t\bar{t}$ and Z+jets. Figures B.5 and B.6 present key kinematic distributions (N-1) in the SR. Where there is a significant expected amount of signal in a given bin the plot is blinded.

Z+jets	430.44 ± 14.29
$t\bar{t}$	128.93 ± 2.84
W+jets	112.36 ± 10.40
Single-top	28.82 ± 1.38
Other	5.97 ± 0.94
DM BBscalar p20 c1	3.79 ± 0.43
DM BBpseudo p20 c1	3.39 ± 0.37
DM BBscalar p200 c1	0.86 ± 0.16
DM BBpseudo p200 c1	0.78 ± 0.13
SM	706.50 ± 17.97

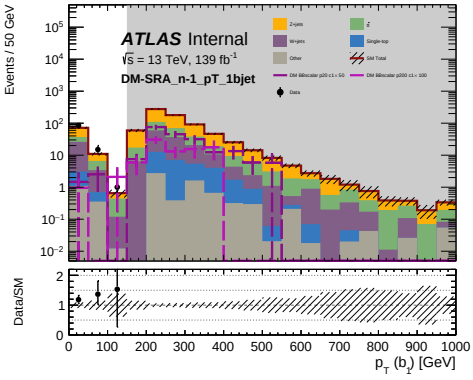
Table B.8 Pre-fit expected yields for DM-SRA.



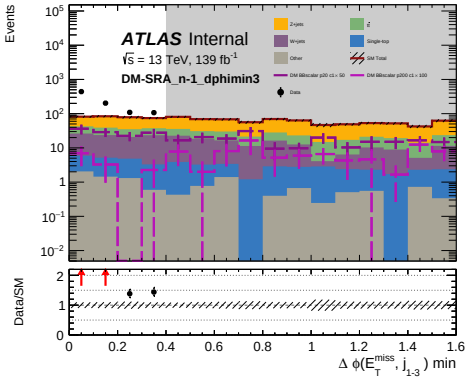
(a)



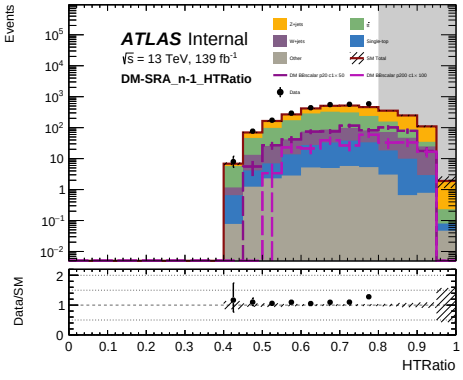
(b)



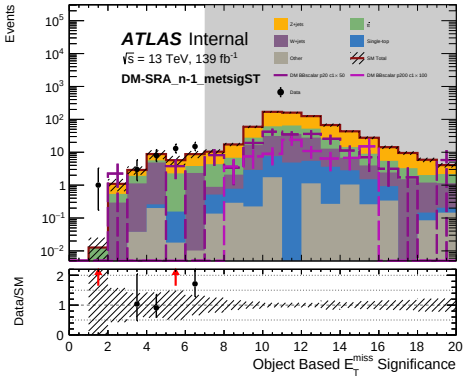
(c)



(d)



(e)



(f)

Fig. B.5 Data vs MC comparisons for DM-SRA for N_j^{good} (a), N_b (b), $p_T(j_1)$ (c), $\Delta\phi(E_T^{\text{miss}}, j^{1-3})$ (d), H_{ratio}^T (e) and S (f).

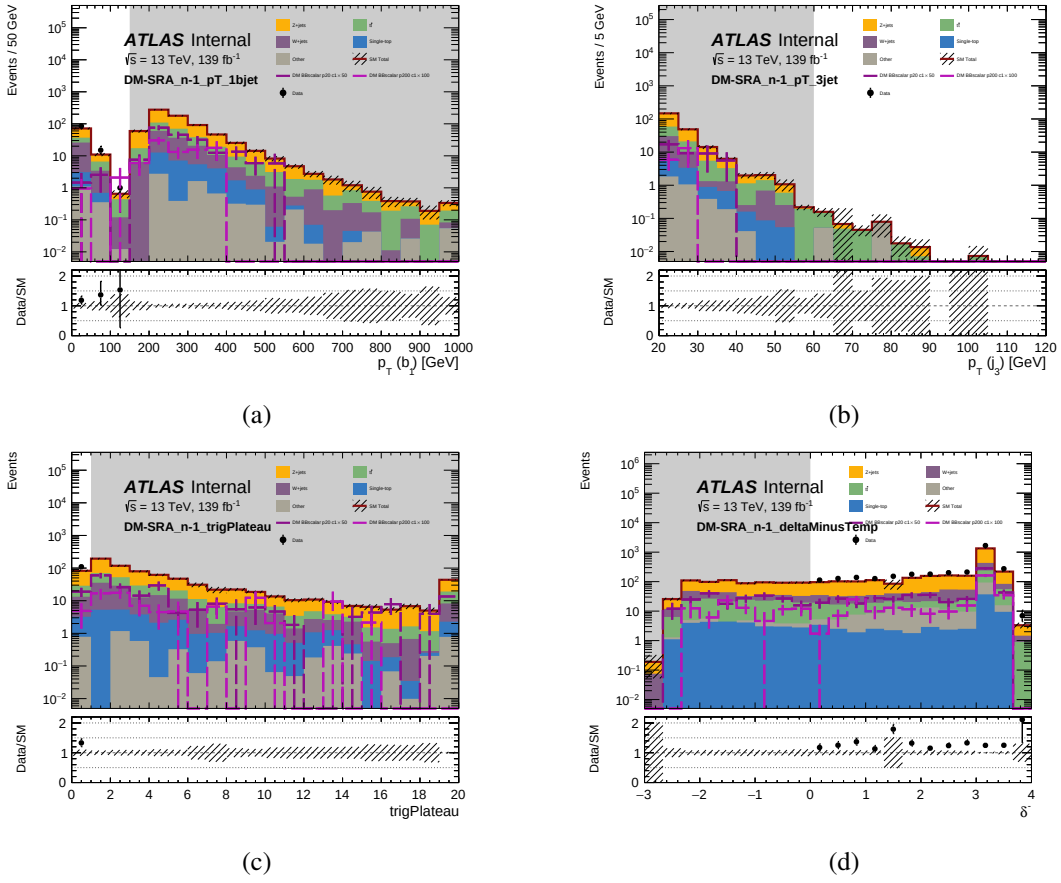


Fig. B.6 Data vs MC comparisons for DM-SRA for $p_T(bj_1)$ (a), $p_T(j_3)$ (b), trigPlateau (c) and δ^- (d).

Appendix C

Systematic Uncertainties

In order for the sensitivity of this analysis to DM signals to correctly be evaluated, it is necessary to properly evaluate the effect of the various sources for these uncertainties. This chapter provides a brief overview of the detector and modelling uncertainties considered in this analysis. These follow the recommendations set out by the ATLAS SUSY working group. The details of the methodology for determining these uncertainties is beyond the scope of this thesis and can be found in reference [97].

C.1 Detector-related Systematic uncertainties

This section will list all of the uncertainties pertaining to the reconstruction of electrons, muons, jets (including heavy flavour jets) and E_T^{miss}. For the final fits in this analysis, these uncertainties are symmetrised and are considered to be gaussian.

C.1.1 Electron related uncertainties

The following uncertainties are used to estimate the up and down variations due to electron resolution, reconstruction and scale factors:

- `EG_RESOLUTION_ALL`, two-sided systematic for the resolution calibration, affecting the shape kinematic distributions
- `EG_SCALE`, two-sided systematic for the energy scale, affecting the shape kinematic distributions
- `EL_EFF_ChargeIDSel_TOTAL`,
  `EL_EFF_ID_TOTAL`, both two-sided systematics for different components of the identification, affecting the normalisation.

- EL_EFF_Iso_TOTAL, two-sided systematic for the isolation scale factor, affecting the normalisation.
- EL_EFF_Reco_TOTAL, two-sided systematic for the reconstruction scale factor, affecting the normalisation.
- EL_EFF_TriggerEff_TOTAL, two-sided systematic for the trigger scale factor, affecting the normalisation.

C.1.2 Jet related uncertainties

The following uncertainties account for the up and down variation inherent in the measurement of jet energy scale (JES):

- JET_GroupedNP_1, first JES NP, two-sided systematic that causes kinematic shape differences.
- JET_GroupedNP_2, second JES NP, two-sided systematic that causes kinematic shape differences.
- JET_GroupedNP_3, third JES NP, two-sided systematic that causes kinematic shape differences.

The following uncertainties account for the up and down variation inherent in the measurement of jet energy resolution (JER):

- JET_JER_EffectiveNP_1 , first JER nuisance parameter (NP), one-sided-symmetrised systematic affecting kinematic shapes.
- JET_JER_EffectiveNP_2 , second JER NP, one-sided-symmetrised systematic affecting kinematic shapes.
- JET_JER_EffectiveNP_3 , third JER NP, one-sided-symmetrised systematic affecting kinematic shapes.
- JET_JER_EffectiveNP_4 , fourth JER NP, one-sided-symmetrised systematic affecting kinematic shapes.
- JET_JER_EffectiveNP_5 , fifth JER NP, one-sided-symmetrised systematic affecting kinematic shapes.
- JET_JER_EffectiveNP_6 , sixth JER NP, one-sided-symmetrised systematic affecting kinematic shapes.

- `JET_JER_EffectiveNP_7restTerm`, seventh, so-called 'rest term', one-sided-symmetrised systematic affecting kinematic. shapes
- `JET_JER_DataVsMC`, `JET_JER_DataVsAFII`, two-sided systematics for jet smearing corrections that are different for full-sim and fast-sim (AFII).

This uncertainty accounts for the jet vertex fraction:

- `JET_JvtEfficiency`, two-sided systematic that effects the normalisation.

C.1.3 Flavour tagging uncertainties

The following uncertainties account for the up and down variation that occurs due to various aspects of flavour tagging:

- `FT_EFF_B_systematics`, two-sided systematic for b-hadron identification that effects the normalisation via the btagging SF.
- `FT_EFF_C_systematics`, two-sided systematic for c-hadron identification that effects the normalisation via the btagging SF.
- `FT_EFF_Light_systematics`, two-sided systematic for light-hadron identification that effects the normalisation via the btagging SF.
- `FT_EFF_extrapolation_systematics`, two-sided systematic for extrapolation to higher p_T for b-tags that effects the normalisation via the btagging SF.
- `FT_EFF_extrapolation_from_charm_systematics`, two-sided systematic for extrapolation to higher p_T for c-hadrons that effects the normalisation via the btagging SF.

C.1.4 E_T^{miss} Uncertainty

The following is a list of the uncertainties that account for the up and down variations arising from the E_T^{miss} soft-term scale and resolution:

- `MET_SoftTrk_ResoPara`, one-sided systematic for the soft-term resolution in the parallel direction to $\mathbf{E}_T^{\text{miss}}$.
- `MET_SoftTrk_ResoPerp`, one-sided systematic for the soft-term resolution in the perpendicular direction to $\mathbf{E}_T^{\text{miss}}$.
- `MET_SoftTrk_Scale`, two-sided for the soft-term resolution scale.

C.1.5 Muon uncertainties

The following is a list of muon related uncertainties:

- `MUON_EFF_BADMUON` , two-sided systematic on the classification of so-called "bad" muons.
- `MUON_EFF_ISO` , two-sided systematic on the isolation efficiency SF.
- `MUON_EFF_RECO` , two-sided systematic on the reconstruction efficiency SF.
- `MUON_SCALE` , two-sided systematic on the muon momentum scale uncertainty.
- `MUON_EFF_Trig` , two-sided systematic for the trigger SF uncertainties.
- `MUON_ID` , `MUON_MS` , two-sided systematics for the resolution uncertainties in the ID and MS systems respectively.
- `MUON_SAGITTA` , two-sided systematic for the uncertainty on the reconstruction of the muon sagitta

C.1.6 Other Uncertainties

Pile-up reweighting uncertainty The uncertainty due to pile-up is calculated called `PRW_DATASF`.

Luminosity uncertainty The overall luminosity uncertainty for Run 2 is taken to be 1.7%. This recommendation comes from the ATLAS data preparation group.

Modelling uncertainties MC modelling uncertainties arise due to choices of various scales (such as factorisation and normalisation) that are used for producing the MC samples. These are generally calculated by varying these scales and symmetrising this uncertainty.

$t\bar{t}$ The hard scatter generation and matching uncertainty is determined by comparing the Powheg+Pythia8 samples used in this analysis with aMcAtNlo+Pythia8. The fragmentation and parton showering uncertainties are determined by comparing the nominal samples with Powheg+Herwig7.

single-top The uncertainties due to the modelling of the single-top process are determined in a similar way to the $t\bar{t}$process. The uncertainty due to the interference of these processes an NLO is determined by comparing the nominal samples to an alternative that uses a different schema.

***V*+Jets** The factorisation, renormalisation, resummation and CKKW matching scale are estimated using a parameterisation of the number of truth jets.